

Grado en Ingeniería en Tecnologías de Telecomunicación
2016-2017

Trabajo Fin de Grado

“Detección del cáncer de piel mediante técnicas de visión artificial”

Paula Fernández Blázquez

Tutor/es

Amaya Jiménez Moreno

Javier López Labraca

Leganés, Septiembre 2017

Agradecimientos

Me gustaría agradecer a mis tutores, Amaya y Javier, por haberme permitido realizar este proyecto que tanto me ha enseñado y por su constante ayuda a lo largo de él.

Gracias a mi familia, en concreto a mis padres Margarita y Antonio, y a mi hermana María, sin vosotros nada de esto hubiera sido posible.

Gracias a mis amigos, con los que he compartido tanto y en los que me he apoyado durante este largo recorrido.

Por último, quiero dar las gracias a mi abuela Amalia y mi madre de nuevo por ser fuente de inspiración en la lucha contra el cáncer.

Resumen

El melanoma no es el cáncer de piel más común, conforma solo el 4% de los cánceres de piel, sin embargo, sí es la forma más peligrosa siendo responsable del 80% de las muertes [3], [4]. En los últimos 25 años, se ha producido un aumento de los casos de este tipo de cáncer de piel, por este y otros motivos, es importante desarrollar herramientas que permitan facilitar el diagnóstico y tratamiento precoces.

El presente Trabajo Fin de Grado se enmarcaría dentro de un sistema CAD, Computer Aided Diagnosis (sistema de ayuda al diagnóstico por computador), y por tanto, su función es la de proporcionar una “segunda opinión” [8], [9], a los dermatólogos en su decisión a la hora de hacer un diagnóstico efectivo a los pacientes. Para llevarlo a cabo nos hemos basado en el artículo “Graph-based Pigment Network Detection in Skin Images”, de M. Sadegui et al. [39], el cual trata de desarrollar un algoritmo para la detección automática de patrones reticulares que están a menudo presentes en los melanomas.

Para la implementación del algoritmo se han utilizado las herramientas de tratamiento de imágenes que el software de Matlab proporciona y un set de 724 imágenes obtenidas del Interactive Atlas of Dermoscopy [49]. A la hora de hacer la clasificación se ha utilizado Python (versión 2.7) y la librería de Machine Learning scikit-learn.

Abstract

Melanoma is not the most common skin cancer as it constitutes only 4% of them; nevertheless, it is the most dangerous form causing 80% of deaths [\[3\]](#), [\[4\]](#). In the last 25 years, there has been an increase in cases of melanoma. For this and other reasons, it is important to develop tools to facilitate early diagnosis and treatment.

The present project would be framed within a CAD system, Computer Aided Diagnosis, and therefore, its aim is to give a “second opinion” [\[8\]](#), [\[9\]](#) to the dermatologist in their decision, at the time to make an effective diagnosis to patients. To achieve this objective, we have based our work in the article “Graph-based Pigment Network Detection in Skin Images” [\[39\]](#), which tries to develop an algorithm for automatic detection of reticular patterns that are often present in melanoma.

For the implementation of the algorithm we have used the image processing tools that Matlab’s software provides and a set of 724 images obtained from the Interactive Atlas of Dermoscopy [\[49\]](#). Python (version 2.7) and the Machine Learning scikit-learn library have been used to classify.

Índice

Lista de Figuras	ix
Lista de Tablas	xi
Lista de Ecuaciones	xii
INTRODUCCIÓN Y MOTIVACIÓN	1
1.1 OBJETIVO PRINCIPAL	2
1.2 INTRODUCCIÓN AL CÁNCER DE PIEL	3
1.3 DETECCIÓN Y TRATAMIENTO DEL MELANOMA.....	6
1.4 MARCO SOCIO-ECONÓMICO Y LEGAL	12
IMAGEN DERMATOSCÓPICA EN LESIONES PIGMENTADAS	14
2.1 PATRONES PRESENTES EN LAS LESIONES DE PIEL PIGMENTADA	16
2.1.1 Patrones globales en lesiones melanocíticas.....	16
2.1.2 Estructuras locales en lesiones melanocíticas	21
ESTADO DEL ARTE	24
3.1 PATRÓN RETICULAR	24
3.2 OTROS PATRONES	26
3.3 MÚLTIPLES PATRONES.....	27
DISEÑO DE LA SOLUCIÓN TÉCNICA.....	29
4.1 CAPTURA DE LAS IMÁGENES.....	30
4.2 PRE-PROCESADO	30
4.2.1 Realce de bordes	30
4.2.2 Eliminación del pelo y las burbujas	31
4.2.3 Detección de bordes	33
4.3 GRAFOS	35
4.3.1 Teoría de grafos.....	35
4.3.2 Búsqueda de ciclos	35
4.4 EXTRACCIÓN DE CARACTERÍSTICAS	40
4.4.1 Características de estructura	42
4.4.2 Características de geometría.....	42
4.4.3 Características de color	43
4.4.4 Características de textura	44

CLASIFICACIÓN Y EVALUACIÓN	48
5.1 MÁQUINA DE VECTORES DE SOPORTE SVM	49
5.2 EVALUACIÓN DEL CLASIFICADOR	51
5.2.1 Matriz de confusión	51
RESULTADOS	53
6.1 DESCRIPCIÓN DEL CONJUNTO DE DATOS	53
6.2 CLASIFICACIÓN DE LAS CARACTERÍSTICAS MEDIANTE UNA SVM LINEAL	54
6.2.1 Clasificación de las lesiones con las 59 características obtenidas	54
6.3 EVALUACIÓN PARA LA CLASIFICACIÓN AUSENTE/ PRESENTE.	55
6.4 RESULTADO DE LA EVALUACIÓN DEL CLASIFICADOR	55
CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO	58
GESTIÓN DEL PROYECTO	60
8.1 PLANIFICACIÓN DEL PROYECTO	60
8.2 PRESUPUESTO	61
BIBLIOGRAFÍA	65
SUMMARY	70

Lista de Figuras

Figura 1. Estructura de la piel. Fuente: http://www.teresewinslow.com/port.asp	4
Figura 2. Etapas del cáncer de piel. Fuente: http://www.skindermatologists.com/melanoma.html	11
Figura 3. Estereomicroscopio	
Figura 4. Dermatoscopio manual	15
Figura 5. Ejemplo patrón reticular	16
Figura 6. Ejemplo patrón globular	17
Figura 7. Ejemplo patrón empedrado	18
Figura 8. Ejemplo patrón homogéneo	18
Figura 9. Ejemplo patrón estallido	19
Figura 10. Ejemplo patrón paralelo	19
Figura 11. Ejemplo patrón multicomponente	20
Figura 12. Ejemplo patrón inespecífico	20
Figura 13. Ejemplo de lesiones con patrón local. (a) Retículo pigmentado típico, (b) Glóbulos y puntos, (c) Hipo pigmentación, (d) Streaks, (e) Velo azul-gris o azul-blancuecino, (f) Estructuras de regresión, (g) Patrón vascular	23
Figura 14. Esquema del sistema	29
Figura 15a. Imagen Original	
Figura 15b. Imagen enfocada	31
Figura 16a. Imagen Original	
Figura 16b. Imagen Enfocada	31
Figura 17. Imagen original enmascarada	
Figura 18. Imagen original enmascarada	32
Figura 19. Imagen original con pelo	32
Figura 20. Resultado de una imagen sin pelo	32
Figura 21. Filtro LoG	33
Figura 22. Imagen original con retículo típico	34
Figura 23. Resultado del filtrado LoG en una imagen con retículo típico	34
Figura 24. Ejemplo de una imagen resultante de aplicar grafos	36
Figura 25. Resultado de aplicar filtro LoG	37
Figura 26. Resultado de aplicar la búsqueda de ciclos mediante grafos	37
Figura 27. Net Mask	
Figura 28. Resultado de aplicar la primera fase	38
Figura 29. Resultado de la fase 1	39
Figura 30. Imagen original en blanco y negro	
Figura 31. Detección de quistes y burbujas	39
Figura 32. Resultado de la fase 1 y 2	39

Lista de Figuras

Figura 33. Imagen original de un patrón reticular típico.....	40
Figura 34. Net mask obtenida tras aplicar la búsqueda de ciclos mediante grafos.....	41
Figura 35. Pigment Network.....	41
Figura 36. Hole mask	41
Figura 37. Representación espacio color HSV. Fuente: http://robocopa.blogspot.com.es/2008/08/bosquejo-en-matlab-del-algoritmo-de.html	43
Figura 38. Proceso de construcción de una matriz GLCM.....	45
Figura 39. Distintas posibilidades para obtener el margen del hiperplano	49
Figura 40. (a) Conjunto de datos original, (b) Datos con separador añadido, (c) Datos transformados	50
Figura 41. Patrón reticular atípico clasificado como típico	56
Figura 42. Diagrama de Gantt	62
Figura 43. Block diagram representation	73

Lista de Tablas

Tabla 1. Análisis de patrones	7
Tabla 2. Sistema de puntuación de la regla ABCD.....	8
Tabla 3. Características método de Menzies.....	9
Tabla 4. Puntuación Regla de los 7 puntos.....	10
Tabla 5. Resultados de la detección de retículo pigmentado	26
Tabla 6. Características de estructura	42
Tabla 7. Matriz de confusión	51
Tabla 8. Tabla de frecuencias de aparición	53
Tabla 9. Resultados clasificación retículo típico o atípico	55
Tabla 10. Resultados clasificación ausente o presente	56
Tabla 11. Coste laboral	61
Tabla 12. Coste de equipamiento.....	63
Tabla 13. Coste total estimado.....	64
Tabla 14. Results of typical or atypical classification	75
Tabla 15. Results of absent or present classification	76
Tabla 16. Estimated budget.....	77

Lista de Ecuaciones

(1) Filtro paso alto bidimensional	30
(2) Filtro LoG para una imagen dada, $f(x, y)$	33
(3) Ratio característica estructura	42
(4) Density Ratio	43
(5) Atypicality.....	44
(6) Entropía	45
(7) Energía.....	46
(8) Contraste	46
(9) Correlación	46
(10) Homogeneidad	47
(11) Conjunto de entrenamiento.....	49
(12) Ecuación hiperplano	49
(13) Distancia perpendicular del hiperplano al origen	49
(14) Función kernel	50
(15) Kernel Lineal	50
(16) Kernel Gaussiano	50
(17) Kernel Polinómico (heterogéneo)	50
(18) Precision	51
(19) Recall	52
(20) Accuracy	52
(21) F1 score	52

Capítulo 1

Introducción y Motivación

Cada año, millones de personas desarrollan cáncer de piel [\[1\]](#). La detección temprana y la cirugía representan el único tratamiento curativo de los pacientes afectados por esta enfermedad. A diferencia de muchos otros tipos de cáncer, el melanoma es visible en la piel; hasta el 70% de los melanomas son identificados por primera vez por el propio paciente (53%) o por familiares cercanos (17%) [\[2\]](#). Aunque no es el cáncer de piel más común, conforma solo el 4% de los cánceres de piel, es responsable del 80% de las muertes [\[3\]](#), [\[4\]](#). Los tumores malignos suponen el otro 96% de los cánceres que pueden aparecer en la piel y se dividen en 2 tipos: carcinoma escamoso, que proviene de la malignización de las células escamosas de la epidermis y representa el 20% de los no melanoma, y carcinoma basocelular, que proviene de la malignización de las células basales de la epidermis. La mayoría de los cánceres de piel de células basales y de células escamosas se deben a la repetida exposición de la piel a los rayos ultravioleta (UV) del sol sin protección, así como a fuentes artificiales, como las camas bronceadoras, las cuales también tienen rayos ultravioleta [\[5\]](#).

La importancia del melanoma reside, más que en su frecuencia, en su rápido aumento en países desarrollados en los últimos 25 años y a que este aumento está relacionado directamente con la exposición solar por motivos estéticos y de ocio. Cada año se diagnostican 36.000 casos nuevos de melanoma en la Unión Europea, que provoca 12.000 defunciones. En España, la incidencia del melanoma es superior 3.600 casos nuevos cada año, siendo éste el tipo de cáncer de piel más trascendente y de mayor mortalidad, y sobre todo, el de más rápida expansión.

Como en el resto de Europa, es un tumor más frecuente entre las mujeres, representando un 2,7% de los cánceres femeninos, mientras que en los hombres solo supone el 1,5%. La incidencia en nuestro país se puede considerar alta, con un ascenso muy importante, especialmente desde los años 90 [6].

Sin embargo, es uno de los cánceres que mejor tratamiento tiene, si se diagnostica a tiempo se pueden llegar a curar el 95% de los casos. El problema es cuando no se detecta a tiempo, lo que puede ocurrir con facilidad ya que no se percibe hasta pasados 20 ó 30 años de la exposición al sol. Esto significa que casi todos los casos que vemos hoy provienen de malos hábitos en los niños y jóvenes de hace más de dos décadas [7].

1.1 Objetivo principal

La detección precoz juega un papel clave en la curación del cáncer de piel, por lo que cobra gran importancia una alternativa de bajo coste que evite la evaluación de cada paciente por parte de los dermatólogos y patólogos. Esta evaluación se realiza a partir de una biopsia que consiste en tomar un trozo de piel para analizarlo después en un microscopio. Realizar una biopsia cuando en realidad no es necesaria puede conducir al paciente a problemas de ansiedad, y existe la posibilidad de que se presente infección o cicatrización queloide (crecimiento exagerado del tejido cicatricial) una vez que ha sido practicada, además del coste para el centro hospitalario que resulta de realizar pruebas que, en realidad, no son necesarias en los casos de lesiones benignas.

El objetivo principal de este Trabajo Fin de Grado consiste en desarrollar un sistema de ayuda al diagnóstico que ayude a los médicos de atención primaria a detectar el cáncer de piel de manera rápida y sencilla utilizando técnicas de visión por ordenador, discriminando las imágenes que presenten melanoma de las que no. Para ello, primero se adecuan las imágenes para poder ser posteriormente procesadas mediante un ordenador, y a continuación, se obtienen características de ellas que nos proporcionan un “resumen” cuantitativo de las propiedades más significativas; finalmente, a partir de esas características se aplicará un clasificador.

Para conseguir este objetivo, se han seguido los siguientes pasos:

- Captura de la imagen de la lesión: la imagen debe estar digitalizada para su análisis por ordenador.

- Pre procesado: es una de las etapas fundamentales del análisis de la imagen. Consiste principalmente en tres pasos que son: enfoque de la imagen, eliminación de elementos que no son de interés como son las burbujas debidas al aceite de inmersión o los pelos de la propia piel y la detección de estructuras de interés.
- Extracción de características: en esta etapa se obtiene una descripción cuantitativa de las estructuras detectadas en el paso previo. Este tipo de descripción será necesario para poder utilizar posteriormente técnicas de aprendizaje automático que nos permitan hacer una clasificación de la lesión.
- Clasificación: a partir de las características obtenidas anteriormente, se aplicarán técnicas de clasificación que nos permitan distinguir el tipo de lesión, o conocer el tipo de estructuras claves presentes en la imagen para ayudar a su diagnóstico.

El presente proyecto se enmarcaría dentro de un sistema CAD, Computer Aided Diagnosis, y por tanto, su función es la de proporcionar una “segunda opinión” [8], [9], a los médicos en su decisión, a la hora de hacer un diagnóstico efectivo a los pacientes. Los sistemas CAD utilizan algoritmos para reconocer patrones en datos de pacientes que indican posibles anomalías. Primero se les entrena a reconocer patrones de una primera base de datos finita con y sin anomalías. Una vez que el sistema ha sido entrenado ya está listo para utilizarse en pacientes nuevos para detectar patrones de anomalías parecidos o descartarlos.

En conclusión, mejoran la exactitud y consistencia del diagnóstico emitido en atención primaria donde los médicos no son especialistas, ayudan a la detección y la clasificación de lesiones a partir de imágenes registradas, pueden incluso evitar pruebas innecesarias con el gasto económico y emocional que pueden suponer para hospital y paciente, y reducen el tiempo de diagnóstico, el cual juega un papel muy importante en la detección del cáncer de piel.

1.2 Introducción al cáncer de piel

Para entender qué es el cáncer de piel debemos saber qué es la piel. Se trata del órgano más grande del ser humano con una extensión de aproximadamente de 1,5 a 2 m², y un peso medio de 3,5 kg lo que representa alrededor del 16% del peso

Capítulo 1. Introducción y Motivación

corporal [10]. Es la primera barrera de protección que nos preserva de las múltiples agresiones externas a las que estamos sometidos en nuestra vida diaria.

La piel sana posee un sistema de reparación muy desarrollado para conseguir mantener la integridad pese a dichas agresiones y caracteriza el aspecto único de los seres humanos. Al mismo tiempo, proporciona al cuerpo un número de funciones fundamentales:

- Es una barrera selectiva, que mantiene un balance de líquidos y electrolitos, e impide la penetración de agentes tóxicos, radiaciones ultravioletas y microorganismos.
- Regula la temperatura corporal, a través de la dilatación o constricción de los vasos sanguíneos, la grasa hipodérmica y la evaporación del sudor.
- Participa en la síntesis de vitamina D.
- Es un órgano de percepción múltiple, a través de millones de terminaciones nerviosas distribuidas en toda su superficie.
- Participa en la vigilancia inmunológica. Dado que sus células sintetizan numerosas sustancias inmunológicamente activas.

Este vital tejido está compuesto por tres capas fundamentales, tal y como se indica en la Figura 1: capa superior o epidermis, capa intermedia o dermis y la capa inferior o hipodermis.

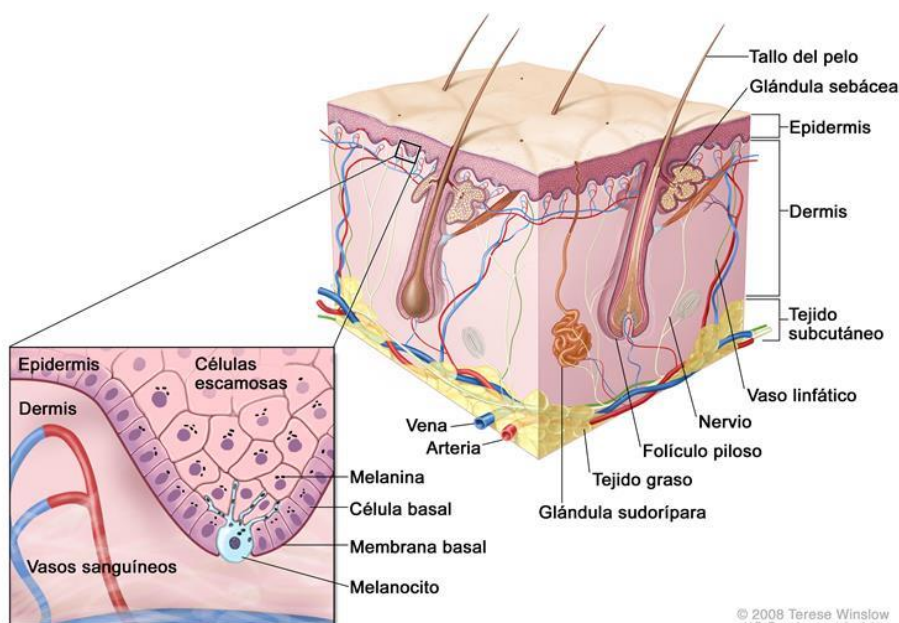


Figura 1. Estructura de la piel. Fuente: <http://www.teresewinslow.com/port.asp>

La hipodermis representa la reserva energética más importante del organismo gracias al almacenamiento y liberación de ácidos grasos. Encima se encuentra la dermis la cual debida a su grosor permite amortiguar los golpes, ayuda a cicatrizar y a curar las heridas gracias a los tejidos conectivos con los que cuenta. Además, nutre la epidermis desde el interior, le aporta resistencia y ayuda a proteger el organismo de posibles infecciones. En la parte superior de la piel se encuentra la epidermis, la capa más delgada y con mayor número de células. Estas células son destruidas y reemplazadas constantemente por la afluencia de nuevas células desde el interior. En la parte más interna de la epidermis se encuentran los melanocitos que son las células encargadas de elaborar el pigmento que da color a la piel, el pelo y los ojos.

Cuando la piel se expone al sol durante un tiempo prolongado, los melanocitos producen más melanina como defensa contra la acción de los rayos ultravioleta, UVA y UVB. En el momento en el que los melanocitos sobrepasan en número a las células normales de la piel aparecen nuevos lunares y pecas. Si bien es cierto, las pecas en su gran mayoría son de origen hereditario. Una persona suele tener entre 10 y 40 y, científicamente, se llaman nevus (del latín naevus, lunar). Estos se clasifican en dos grupos: nevus congénitos, que pueden ser de nacimiento o aparecer con los años; y nevus adquiridos, que se desarrollan a partir del año de nacimiento y pueden ser típicos o atípicos [11].

Es la epidermis donde se van a producir las lesiones que pueden degenerar en cáncer de piel, melanoma y no melanoma. La principal diferencia es que el no melanoma se genera a partir de células de la piel distintas a los melanocitos, como son las células basales y las células escamosas. Ambos tipos de cáncer de piel pueden ocurrir en cualquier parte del cuerpo, aunque es bastante frecuente encontrarlos en la cabeza, cara, cuello, dorso de las manos brazos y piernas, ya que estas zonas están más expuestas a los rayos del sol. En los hombres, la incidencia es mayor entre los 50 y 59 años y su localización es más frecuente en la espalda. En las mujeres es más frecuente entre los 40 y 49 años, y suele presentarse en las piernas [2]. Aun así, el cáncer de piel puede afectar a cualquier persona, aunque las personas con ciertos factores de riesgo tienen más probabilidad que otras de padecerlo.

- Las personas de raza blanca tienen mayor riesgo de padecer cáncer de piel que las personas de raza negra o los hispanos. Esto se debe a que la melanina tiene un efecto protector en las personas con piel más oscura. Por ejemplo, en un afroamericano de tez oscura la melanina le da un factor de protección solar (FPS) de aproximadamente 13.4, comparado con el 3.4 de una piel clara [12].

- Las personas de edades avanzadas. Probablemente esto se debe a la acumulación de la exposición solar con el transcurso del tiempo. Se estima que 1 de cada 5 personas de 70 años va a padecer al menos un cáncer de piel [4].
- Aquellas con antecedentes familiares de primer grado (madre, padre, hermano/a, hijo/a) que han tenido cáncer de piel, melanoma. El aumento en el riesgo pudiera deberse a que compartieron un estilo de vida de exposición frecuente, mismo tono de piel, ciertos cambios genéticos (mutaciones), o una combinación de factores.
- Una persona que ha padecido melanoma tiene un mayor riesgo de padecer melanoma nuevamente, el riesgo relativo de desarrollar un segundo melanoma es 70 veces superior al de desarrollar un primer melanoma [4].
- Las personas cuyo sistema inmunológico se ha debilitado, debido a ciertas enfermedades o tratamientos médicos, tienen más probabilidades de padecer cáncer de piel.
- Existe una correlación entre el riesgo de melanoma y la latitud [13]. También influye el período de tiempo durante el que ha vivido una persona en países de latitudes ecuatorianas, lo que significa una exposición excesiva al sol mantenida durante un período de la vida.
- La exposición a carcinógenos, los traumatismos o cicatrices, las lesiones por radiación crónica y las infecciones virales, son algunos de los factores que predisponen a la aparición del cáncer de piel.

1.3 Detección y tratamiento del melanoma

El melanoma maligno de piel ocupa la séptima posición entre los tumores más prevalentes en la población [14], por lo que se debe tener más conocimiento sobre la prevención y la identificación. Por lo general, el melanoma puede detectarse en sus inicios. Cada persona puede desempeñar una función importante en la detección temprana del cáncer de piel, cuando es más probable que sea curable.

Uno de los primeros signos de alarma para detectar este tipo de enfermedad es una herida que no cicatriza o la aparición de nuevas manchas en la piel. Por lo que es muy

importante para detectar a tiempo el posible melanoma la autoexploración. Los dermatólogos recomiendan la autoexploración periódica en busca de cualquier lunar, llaga, protuberancia, imperfección, marca o cambio inusual en el aspecto o la sensación de un área de la piel.

Para evaluar si una lesión melanocítica es benigna o es un melanoma el médico puede usar distintos métodos. Hay muchas reglas clínicas para diagnosticar la presencia de melanoma que han ido surgiendo conforme se ha profundizado en la materia.

El análisis de patrones es el método más usada por los dermatólogos y con la que mejores resultados se obtienen; aunque por otra parte, es el que requiere de mayor experiencia. Existen dos modalidades en función de si se detectan patrones globales o locales. Los patrones son las estructuras que cubren la mayor parte de la lesión permitiendo una categorización preliminar rápida. Por otra parte, las estructuras locales son aquellas presentes en regiones concretas de la lesión. En el siguiente capítulo se expone con mayor profundidad, explicándose con mayor detalle cada uno de los patrones.

La Tabla 1 resume los patrones locales y globales que se tienen en cuenta en el análisis de patrones.

Patrones globales	Patrones locales
Reticular Globular Empedrado Homogéneo Estrellado Paralelo Multicomponente Inespecífico	Retículo pigmentado Glóbulos y puntos Áreas de hipo pigmentación Streaks Velo azul-gris o azul-blancuecino Estructuras e regresión Patrón vascular

Tabla 1. Análisis de patrones

La regla ABCD fue el primer método alternativo al “análisis de patrones” para el diagnóstico dermatoscópico del melanoma. Fue desarrollado por Stolz et al. [15] en 1994 y pretendía facilitar el diagnóstico del melanoma a aquellos observadores no expertos en dermatoscopia mediante una simplificación del método análisis de patrones. Su nombre hace referencia a las cuatro características que son evaluadas: Asimetría, Borde, Color y Diferencias estructurales. A partir de la observación de dichas

Capítulo 1. Introducción y Motivación

características se asigna a la lesión una puntuación semi-cuantitativa que permite determinar con precisión razonable si se trata de un melanoma o de una lesión benigna.

- A: Lesión asimétrica. Se divide la lesión en dos ejes de 90° consiguiendo la mayor simetría posible, y se valora la asimetría con respecto al color, forma y estructuras en ambos lados de los ejes. Se otorga una puntuación de 0 si no existe asimetría en ningún eje, 1 si se encuentra asimetría respecto de un único eje, y 2 si es asimétrica respecto de ambos ejes. De modo que una lesión puede tener una puntuación de 0 a 2.
- B: Bordes irregulares. La lesión se divide en 8 segmentos y se puntúa con 1 cada porción que presente una terminación abrupta del borde. De manera que la puntuación mínima que se puede obtener es 0 y la máxima 8.
- C: Distintos colores en la misma lesión. La puntuación de color es la suma de los diferentes colores encontrados en la lesión. Se valora la presencia de 6 colores: blanco, marrón claro, marrón oscuro, azul-gris, rojo y negro. El blanco solo puntuará si es más claro que la piel adyacente.
- D: Diferencias estructurales. Se suma un punto por cada patrón estructural presente en la lesión. Las posibilidades son: red pigmentada, áreas homogéneas, manchas, glóbulos, y streaks. La puntuación máxima es 5 y la mínima es 1.

La Tabla 2 describe el sistema de puntuación empleado por la regla ABCD.

Criterio	Rango de puntuación	Factor	Puntuación máxima/mínima
Asimetría	0-2	1.3	0.0-2.6
Bordes	0-8	0.1	0.0-0.8
Color	1-6	0.5	0.5-3.0
Diferencias estructurales	1-5	0.5	0.5-2.5

Tabla 2. Sistema de puntuación de la regla ABCD

En función de la puntuación **P** calculada para una lesión se emite un diagnóstico basado en los siguientes rangos:

- $P < 4,75 \rightarrow$ Lesión benigna
- $4,75 < P < 5,45 \rightarrow$ Lesión sospechosa
- $P > 5,45 \rightarrow$ Lesión maligna

Dos años más tarde, en 1996, Menzies publicó el método que lleva su nombre, método de Menzies [16]. Este método valora 11 criterios dermatoscópicos; estos se dividen en criterios negativos, que no deben estar presentes para el diagnóstico de melanoma, y criterios positivos, es decir, que alguno de ellos debe cumplirse para el diagnóstico de malignidad. Las características mencionadas se muestran en la siguiente tabla.

Características benignas (negativas)	Características malignas (positivos)
<p>Monocromía</p> <p>Simetría</p>	<p>Patrones asimétricos</p> <p>Pigmento reticular</p> <p>Colores múltiples</p> <p>Velo azul-blanquecino</p> <p>Puntos/glóbulos oscuros</p> <p>Pseudópodos</p> <p>Distribución radial</p> <p>Zonas sin pigmento</p> <p>Puntos negros en la periferia</p>

Tabla 3. Características método de Menzies

Más tarde, en 1998, Argenziano et al presentó la Regla de los 7 puntos [17]. Al igual que el método de Menzies, este método pretende simplificar el diagnóstico a través de la identificación de ciertos patrones en la lesión. En este caso se puntúa cada característica en función de su relevancia en las lesiones malignas. Existen tres criterios mayores (retículo pigmentado atípico, velo azul-blanquecino y patrón vascular atípico) que recibirán una puntuación de 2 y cuatro criterios menores (líneas irregulares, pigmentación irregular, puntos irregulares y áreas de regresión) que tendrán una puntuación de 1. El diagnóstico es positivo (melanoma) en caso de que la suma de las puntuaciones de los patrones encontrados es igual o mayor que 3.

La siguiente tabla presenta las características buscadas y sus puntuaciones.

Criterio	Características	Puntuación
Mayor	Pigmento reticular atípico Patrón vascular atípico Velo azul-blanquecino	2
Menor	Líneas irregulares Pigmentación irregular Puntos irregulares Áreas de regresión	1

Tabla 4. Puntuación Regla de los 7 puntos

Una vez que el especialista sospecha de la existencia de alguna anomalía, realizará una prueba más exacta, una biopsia. Esta prueba consiste en la obtención de tejidos u otros materiales procedentes del organismo vivo, para su examen microscópico con fines diagnósticos [18]. Existen varias formas para realizar una biopsia de piel las cuales dependerán de la zona, tamaño y tipo de lesión.

- Biopsia por raspado.
- Biopsia por sacabocados.
- Biopsia por escisión.

La biopsia por raspado es la menos invasiva de las tres técnicas y a menudo se hace cuando se sospecha de cáncer de piel. El médico extrae las capas más superficiales de la piel. La biopsia en sacabocados se emplea con mayor frecuencia para lesiones cutáneas más profundas. El médico extrae un pequeño pedazo de piel, utilizando un instrumento filoso y hueco. Finalmente, una biopsia por escisión se lleva a cabo para extraer toda la lesión y se realiza cuando el medico sospecha de la existencia de melanoma. Se inyecta un anestésico en el área; luego, se extrae toda la protuberancia, mancha o úlcera, profundizando hasta donde sea necesario para obtener toda el área afectada. Se cierra luego la incisión con puntos de sutura y se aplica presión en el área para detener cualquier sangrado.

Los resultados se analizan con un microscopio. Si la lesión de piel es benigna, es posible que no se necesite ningún tratamiento adicional, sin embargo, si es maligna

entonces será necesario saber en qué etapa de las 5 mostradas en la Figura 2 se encuentra [19], [20].

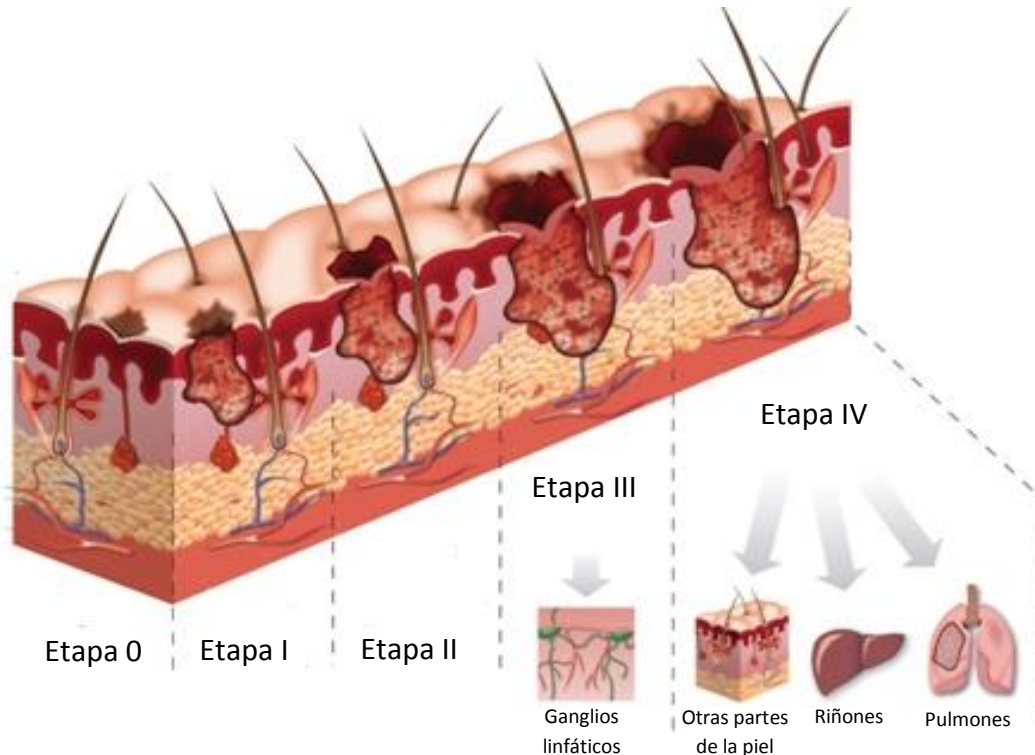


Figura 2. Etapas del cáncer de piel. Fuente: <http://www.skindermatologists.com/melanoma.html>

En la etapa 0 los melanomas no han crecido más allá de la epidermis. Por lo general, estos melanomas son tratados con cirugía y si se descubre que los bordes de la muestra que se extrajo contienen células cancerosas, es posible que se repita una escisión del área. La tasa de supervivencia a 5 años es del 97%.

El melanoma en la etapa I tiene un grosor menor de 2mm y se encuentra en la epidermis, pudiendo llegar a la capa superior de la dermis, sin haberse diseminado a los ganglios linfáticos cercanos. Estos ganglios son parte del sistema linfático, una red de órganos, conductos y vasos que apoyan al sistema inmunitario del cuerpo. El cáncer puede propagarse a través de ellos produciendo metástasis. La tasa de supervivencia a 5 años es del 90-95%.

En la etapa II el cáncer se ha diseminado a la dermis, pero no más allá de ésta ni a los ganglios linfáticos y tiene un grosor de 2 y 4 mm. La tasa de supervivencia a 5 años es del 45-85%. Este amplio rango de supervivencia refleja la variación en la evaluación de los ganglios linfáticos.

Ya es en la etapa III cuando el tumor, de 4 mm, afecta a los tejidos situados debajo de la piel y a los ganglios linfáticos cercanos a la lesión principal, por lo que es necesaria una linfadenectomía o cirugía para extirpar los ganglios linfáticos afectados. Además, junto con la escisión se recomienda un tratamiento adicional. La tasa de supervivencia a 5 años es de 45%.

La etapa IV es la fase más avanzada de melanoma. El tumor afecta a los ganglios linfáticos alejados del tumor original o a órganos, por lo que es difícil de curar. La tasa de supervivencia a 5 años es del 10%.

En ciertas ocasiones, el melanoma podría regresar al mismo lugar donde se originó el tumor inicial, algunas veces incluso en la cicatriz de la cirugía, o en partes distintas del cuerpo. El tratamiento del melanoma que reaparece después del tratamiento inicial depende de la etapa del melanoma original, del tipo de tratamiento inicial y del lugar de la recurrencia.

Cuando el melanoma está en las primeras etapas (0, I y II) es muy tratable a través de procedimientos quirúrgicos que eliminan las células cancerosas. A medida que el cáncer empeora y se propaga (III y IV), el melanoma puede ser difícil de tratar con las terapias convencionales como la cirugía, la quimioterapia y la radioterapia. Por ello, es tan importante el diagnóstico de forma rápida, el cual se puede conseguir mediante la detección automática sin la necesidad de realizar una biopsia al paciente. Con este sistema se consigue además que los dermatólogos puedan controlar de un modo muy preciso la evolución de los lunares y otras lesiones sospechosas, incluso antes de que degeneren en un cáncer de piel.

1.4 Marco socio-económico y legal

En lo que se refiere al marco legal se debe tener en cuenta que para llevar a cabo investigaciones con imágenes médicas de pacientes previamente se debe pedir el consentimiento de estos. Desde el punto de vista de las normativas internacionales, el párrafo 23 de la Declaración de Helsinki de la Asociación Médica Mundial¹, establece que “deben tomarse toda clase de precauciones para resguardar la intimidad de la persona que participa en la investigación y la confidencialidad de su información personal y para reducir al mínimo las consecuencias de la investigación sobre su integridad física, mental y social”. El párrafo 25 a su vez establece que “para la investigación médica en que se utilice material o datos humanos identificables, el

¹ Declaración de Helsinki de la Asociación Médica Mundial, octubre 2008.

médico debe pedir normalmente el consentimiento para la recolección, análisis, almacenamiento y reutilización.”

De forma similar, nuestra normativa nacional parece decantarse por que la confidencialidad de la información de los pacientes no puede quebrantarse sin su consentimiento. La Ley General de Sanidad², determina que los ciudadanos tienen derecho a la confidencialidad de toda la información relacionada con su proceso y con su estancia en instituciones sanitarias públicas y privadas que colaboren con el sistema público. También la Ley de autonomía del paciente³, hace referencia a la protección de la intimidad señalando que toda persona tiene derecho a que se respete el carácter confidencial de los datos referentes a su salud y a que nadie pueda acceder a ellos sin previa autorización amparada por la ley.

Nuestras imágenes han sido tomadas del Interactive Atlas of Dermoscopy [49] el cual cumple con la normativa expuesta anteriormente.

Económicamente, mediante este método que no requiere de cirugía se puede detectar si una lesión es benigna o maligna, reduciendo así los costes materiales y de personal hospitalario, realizando únicamente las pruebas que sean necesarias para las lesiones malignas que lo requieren, lo que es clave para el adecuado funcionamiento del sistema de salud.

La realización de una biopsia puede provocar sangrado, infección, problemas de cicatrización y lesiones accidentales en las estructuras adyacentes. Es por ello, que es mejor evitarla para no provocar al paciente alguno de estos daños.

² Ley 14/1986, de 25 de abril, General de Sanidad

³ Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica

Capítulo 2

Imagen dermatoscópica en lesiones pigmentadas

La dermatoscopia, también llamada microscopía de epiluminiscencia, es una técnica no invasiva que mejora el diagnóstico clínico de las lesiones cutáneas, especialmente las pigmentadas [21], y permite identificar mediante un instrumento óptico, llamado dermatoscopio, estructuras de la piel más profundas que no son apreciables a simple vista por el ojo humano. La dermoscopia digital es considerada como uno de las técnicas más rentables para identificar y clasificar el cáncer de piel como el melanoma [22]. Este método permite almacenar imágenes en un ordenador gracias a un software, por medio de una cámara de vídeo que toma mapas corporales, imágenes macroscópicas y microscópicas. Esta técnica también permite digitalizar las imágenes y observar los cambios en el tiempo de las lesiones melanocíticas.

Sus orígenes se remontan al año 1655, cuando el médico Pierre Borrel estudió con una lente los capilares de la uña y relacionó los cambios observados con distintas enfermedades. Dos siglos más tarde, en 1878, Abbe sugirió las ventajas de la aplicación de aceite de inmersión en lugar de agua para mejorar la imagen tomada, ventaja que utilizó Unna en 1893 al reconocer que las capas superiores de la epidermis bloqueaban la penetración de la luz. Decidió entonces aplicar la idea de Abbe utilizando aceites solubles y otros fluidos sobre la piel para hacerla más traslúcida; de esa manera estudió el lupus vulgar al colocar una lente de vidrio directamente sobre la piel de un paciente. A este procedimiento lo llamó diascopia. Con el paso del tiempo la técnica se perfeccionó y se implantó en la dermatología. Saphier fue el primer autor que acuñó el término de “Dermatoscopia” con las connotaciones actuales y el que comenzó a

Capítulo 2. Imagen dermatoscópica en lesiones pigmentadas

evaluar nevus melanocíticos, describiendo por primera vez los glóbulos, parámetro dermatoscópico actualmente vigente.

Posteriormente en 1971, tras unos 20 años de ausencia de referencias a la dermatoscopia, Rona Mackie la devolvió a la actualidad para destacar, por primera vez, su utilidad en el incremento de la precisión diagnóstica preoperatoria de las lesiones pigmentadas y en la diferenciación entre las lesiones benignas y malignas. Esta línea de investigación fue seguida principalmente en Europa por grupos austriacos y alemanes a partir de la siguiente década. En 1981 Fritsch y Pechlaner, introdujeron el estereomicroscopio [15], representado en la Figura 3. Este aparato consta de dos microscopios completos, cada uno con su objetivo y ocular. Al no coincidir sus ejes ópticos, las imágenes formadas en los oculares son distintas, por lo que se pueden observar imágenes en tres dimensiones. Sin embargo, este instrumento obstaculizaba la aplicación clínica por ser muy pesado, caro e incómodo. Por lo que la técnica tenía un procedimiento difícil y consumía mucho tiempo [15].

En 1990, Braun-Falco et al. [23] diseñaron el dermatoscopio manual representado en la Figura 4. Es un instrumento pequeño, manejable, económico y fácil de utilizar que emplea un sistema de magnificación con una fuente de luz que ilumina la piel y permite un aumento de 10 a 400 veces.



Figura 3. Estereomicroscopio
Fuente: <http://www.equiposylaboratorio.com>



Figura 4. Dermatoscopio manual
Fuente: <https://galeriamedica.mx>

Gracias a la dermatoscopia se pueden determinar las características de la lesión e indicar un tratamiento o la necesidad de alguna otra exploración complementaria más compleja como la biopsia. Cuando el dispositivo de dermatoscopia se acopla a un sistema informático se denominada microscopía de epiluminiscencia digitalizada [24]. Esta técnica permite estudiar las estructuras de la piel y los patrones sin necesidad de extraer muestras de la piel del paciente. Utilizando este nuevo método se puede visualizar una gama completamente nueva de características visuales como variación

de color, asimetría, características de textura y geometría que pueden ser de importante relevancia en el proceso de diagnóstico.

2.1 Patrones presentes en las lesiones de piel pigmentada

La presencia de ciertos patrones dermatoscópicos orienta a los dermatólogos en la detección precoz de melanoma, que en ocasiones puede confundirse clínicamente con una queratosis seborreica pigmentada, un carcinoma basocelular pigmentado, un hemangioma trombosado, un nevo de Spitz o un nevo azul. Por este motivo es importante conocer los patrones que sugieren la presencia de un melanoma.

2.1.1 Patrones globales en lesiones melanocíticas

En este apartado se describen los principales tipos de patrones globales que pueden estar presentes en el melanoma de forma predominante [\[25\]](#), [\[26\]](#).

- Patrón reticular

Los nevus de patrón reticular son más frecuentes en adultos y se distribuyen en el tronco [\[27\]](#), [\[28\]](#). Se trata de retículos pigmentados que cubren la mayoría de la lesión y pueden encontrarse tanto en las lesiones benignas como en las malignas. Poseen estructuras lineales de color marrón-negro. Mostramos un ejemplo en la Figura 5.



Figura 5. Ejemplo patrón reticular

Capítulo 2. Imagen dermatoscópica en lesiones pigmentadas

Los patrones reticulares que se presentan como retículos regulares, con líneas delgadas y con un borde que se difumina en la periferia o bien poseen una distribución simétrica son indicativos de que la lesión es benigna. Al contrario, si los retículos son prominentes, irregulares, con mallas de diferentes tamaños y configuración o con forma abrupta en su periferia, son indicativos de melanoma.

- Patrón globular

Los nevus de patrón globular predominan en la infancia y adolescencia, y suelen localizarse en la cara, el cuello y en la parte superior del tronco [27], [28]. Consisten en estructuras redondas u ovales, numerosas, de diferente tamaño y color marrón y/o negro grisáceo. Un ejemplo se presenta en la Figura 6.

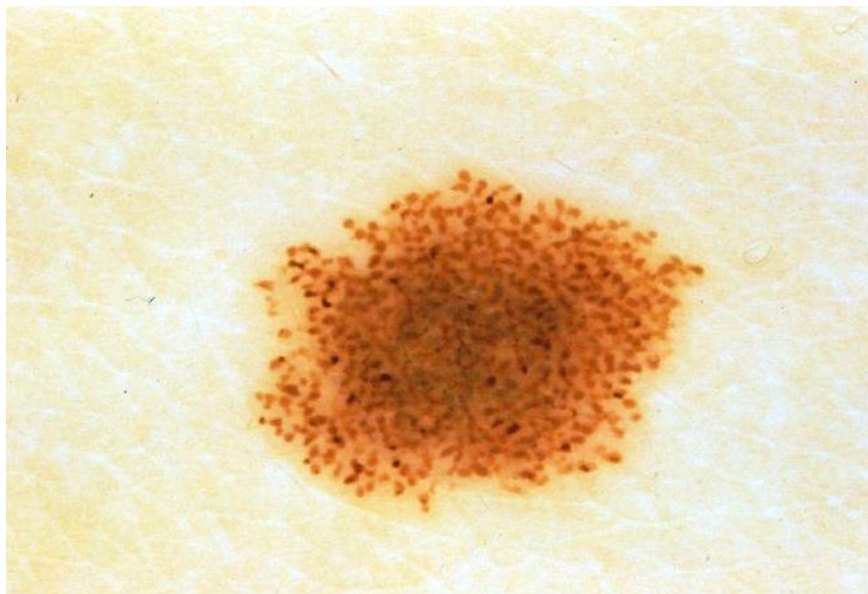


Figura 6. Ejemplo patrón globular

Cuando los glóbulos son regulares y uniformes en tamaño y de forma simétrica y homogénea, entonces representa una lesión benigna. En los melanomas este patrón presenta diferentes tamaños, formas y colores además de estar distribuidos de modo irregular.

- Patrón empedrado

Son glóbulos grandes, distribuidos de forma muy cercana, en forma poligonal semejando un empedrado o mosaico. Este tipo de patrón, del que mostramos una

imagen en la Figura 7, si aparece distribuido de forma simétrica quiere decir que se trata de una lesión benigna. Por el contrario, si no son uniformes y en ocasiones se presentan de color rojo entonces indica que es una lesión maligna.

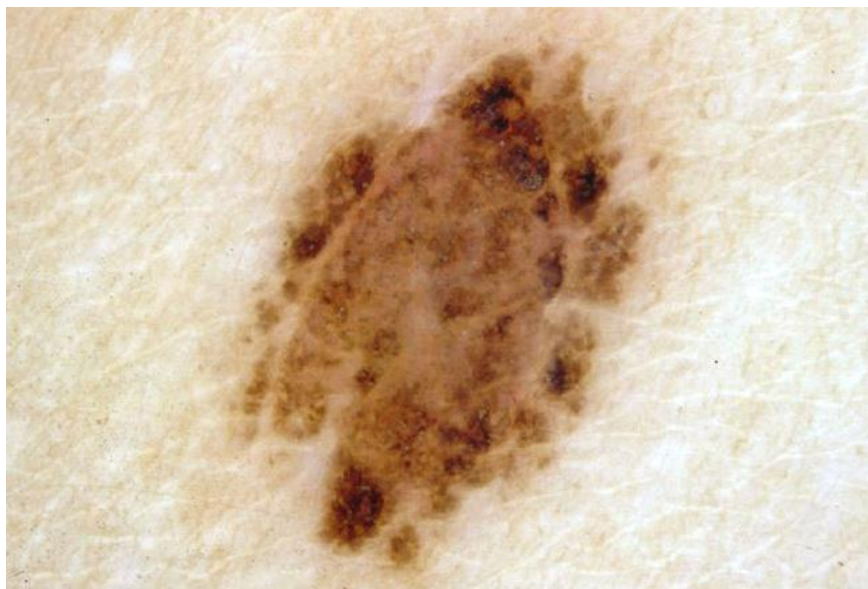


Figura 7. Ejemplo patrón empedrado

- Patrón homogéneo

La lesión presenta una superficie homogénea y no tiene otras características locales distintivas. La pigmentación es difusa de color marrón o en tonos entre gris azulado y gris negruzco. Mostramos un ejemplo de este patrón en la Figura 8.



Figura 8. Ejemplo patrón homogéneo

- Patrón estallido

Se caracteriza por la presencia de proyecciones que adoptan una distribución radial y regular en toda la periferia de la lesión pigmentada, como se puede apreciar en la Figura 9. Como en el caso del patrón empedrado este tipo de patrón se presenta en lesiones benignas con una distribución simétrica y cuando su distribución es irregular o asimétrica la lesión es más sospechosa de ser maligna.

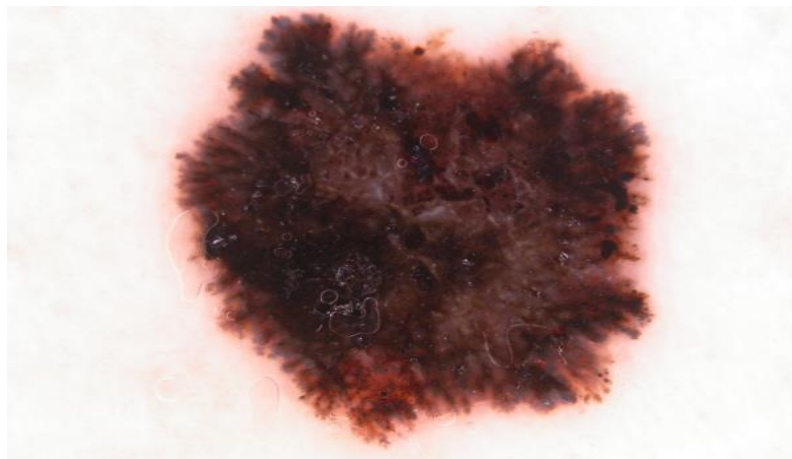


Figura 9. Ejemplo patrón estallido

- Patrón paralelo

La pigmentación de la piel en este patrón sigue los surcos o las crestas de los dibujos formados por los pliegues cutáneos presentes en las palmas de las manos o plantas de los pies. Las líneas paralelas suelen ser delgadas y homogéneas y el pigmento sigue los surcos cutáneos. De ocurrir esto la lesión es benigna. A continuación, mostramos en la Figura 10 un ejemplo.



Figura 10. Ejemplo patrón paralelo

- Patrón multicomponente

Es la combinación de 3 o más estructuras dermatoscópicas distintivas (retículo pigmentado, glóbulos, velo blanquecino...), como mostramos en la Figura 11. Si las estructuras individuales aparecen de forma homogénea, la lesión será benigna. Por el contrario, serán malignas aquellas cuyas estructuras individuales sean atípicas, irregulares y/o asimétricas.

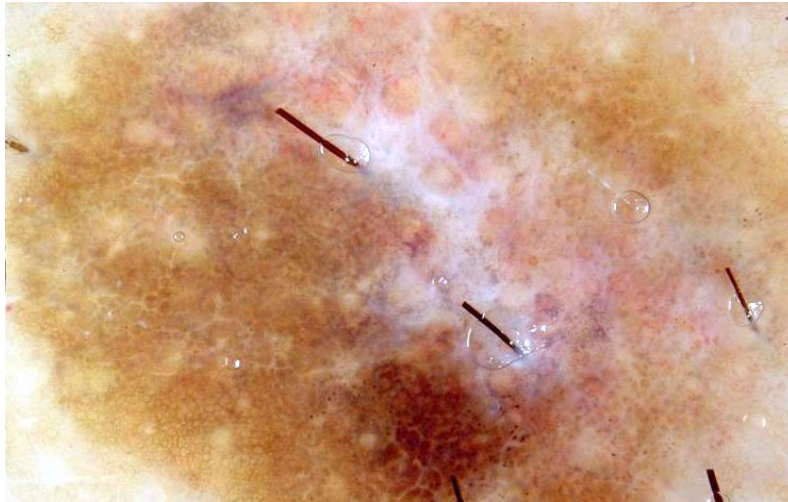


Figura 11. Ejemplo patrón multicomponente

- Patrón inespecífico

Lesión pigmentada sin criterios suficientes para catalogar su patrón como uno de los mencionados anteriormente. Mostramos un ejemplo en la Figura 12.

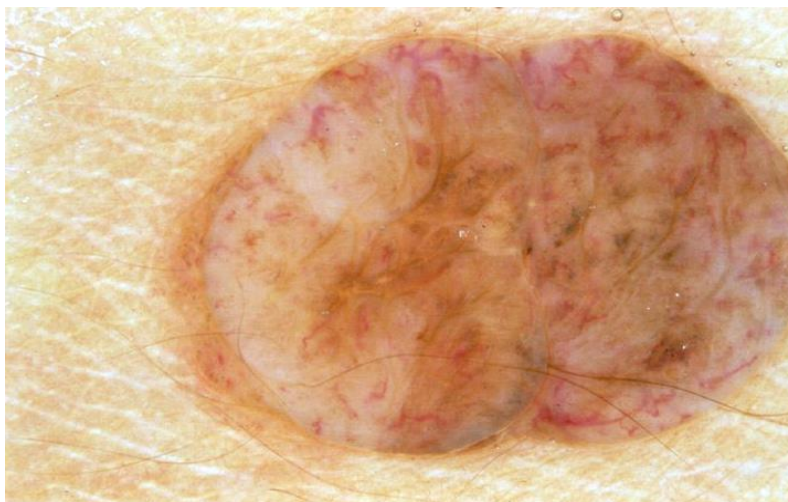


Figura 12. Ejemplo patrón inespecífico

2.1.2 Estructuras locales en lesiones melanocíticas

Cuando existe presencia de características dermatoscópicas específicas en diferentes regiones de la misma lesión, estos patrones son llamados locales [25], [26]. Son estructuras dermoscópicas como el retículo pigmentado, glóbulos y puntos, áreas de hipo pigmentación, streaks, velo azul-blanquecino o azul-gris, estructuras de regresión y patrones vasculares cuya descripción de apariencia se explica a continuación.

- Retículo pigmentado

Se trata del patrón local más estudiado y es el que se va a tratar en este proyecto. La razón es que es el más común de los patrones locales en los melanomas. Una red de pigmentos puede ser típica o atípica, lo que es determinante a la hora de valorar el riesgo de una lesión. Será típica si el patrón presenta una distribución regular de color marrón y con los retículos estrechamente espaciados. Será atípica si la distribución es irregular de color negra, gris o marrón.

- Glóbulos y puntos

Son estructuras de color negro o marrón con forma circular u oval de tamaño mayor a 0,1 mm, distribuidas de forma regular en el caso de las lesiones benignas y de forma irregular en las malignas.

- Hipo pigmentación

Se produce por reducción en la cantidad de melanina. Cuando está presente en el melanoma se presenta en forma focal, multifocal o difusa.

- Streaks

Son estructuras lineales radiadas que se localizan en la periferia de la lesión. Si se distribuyen de forma irregular y asimétrica son altamente sugestivas de la fase de extensión superficial del melanoma.

- Velo azul-gris o azul-blانquecino

Es uno de los parámetros más específicos de melanoma. Se trata de un área irregular, desestructurada que suele estar recubierta por un velo blanquecino que le otorga una apariencia de vidrio esmerilado. La pigmentación no abarca la totalidad de la lesión y en general, corresponde a un área clínicamente elevada de la lesión.

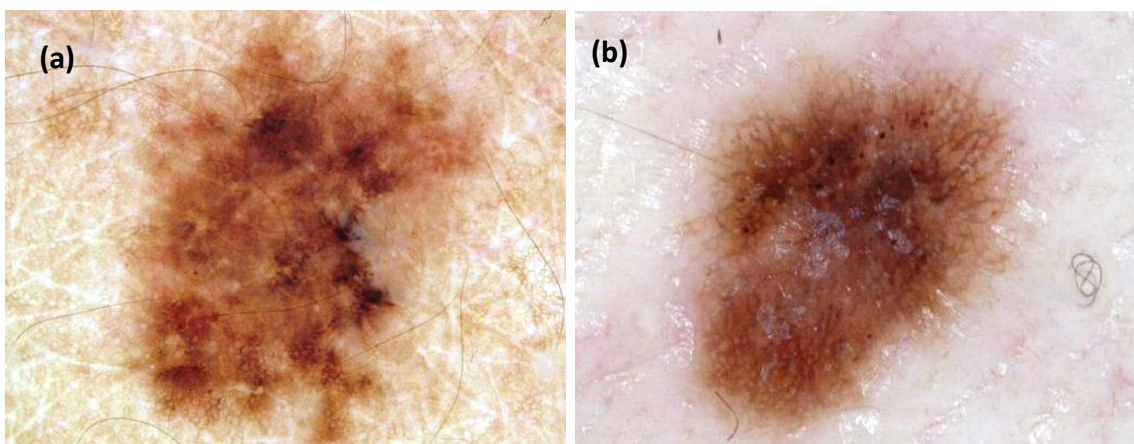
- Estructuras de regresión

Pueden distinguirse áreas de regresión azul, blanca y mixta. Las áreas de regresión azul son denominadas “áreas con punteado azul” o “en pimienta” y se caracterizan por un color azul-grisáceo con un punteado fino en su interior. Las áreas de regresión blancas, también llamadas pseudocicatriciales, recuerdan una cicatriz superficial. Las áreas de regresión mixtas contienen ambos tipos.

- Patrón vascular

Son vasos sanguíneos de distribución irregular, de apariencia lineal, puntiforme o en “horquilla”, y/o eritema dentro de las estructuras de regresión.

A continuación mostramos en la Figura 13, de izquierda a derecha, un ejemplo de cada uno de los patrones locales mencionados anteriormente.



Capítulo 2. Imagen dermatoscópica en lesiones pigmentadas

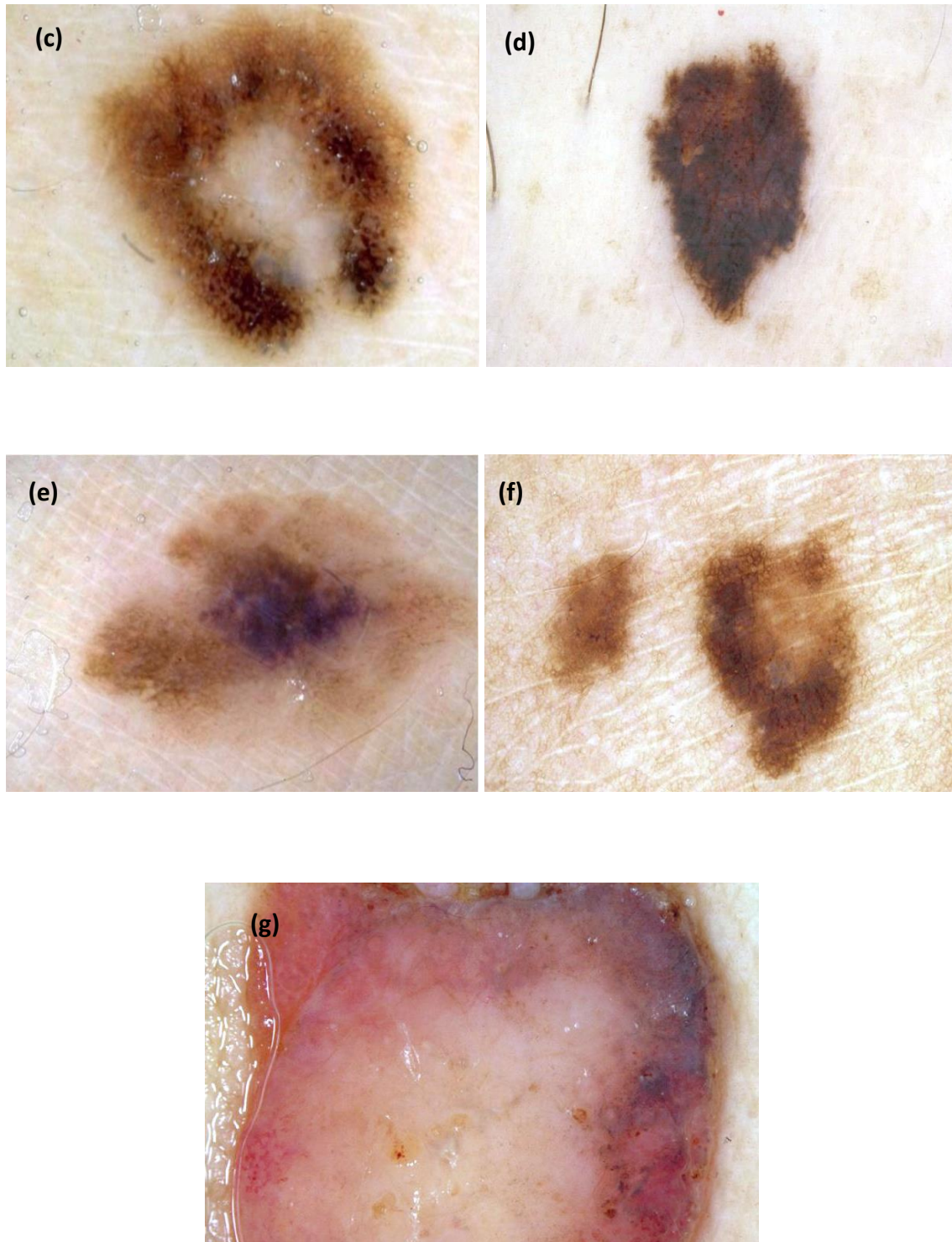


Figura 13. Ejemplo de lesiones con patrón local. (a) Retículo pigmentado típico, (b) Glóbulos y puntos, (c) Hipo pigmentación, (d) Streaks, (e) Velo azul-gris o azul-blانquecino, (f) Estructuras de regresión, (g) Patrón vascular.

Capítulo 3

Estado del arte

Existen numerosas propuestas de sistemas de asistencia al diagnóstico [29] para lesiones en la piel basados en diferentes criterios de reconocimiento de patrones. Algunas de ellas intentan distinguir un único patrón en la lesión pigmentada y otras intentan buscar varios patrones.

3.1 Patrón reticular

La detección automática de patrones reticulares ha recibido mucha atención en los últimos años. Fleming et al. [30] propusieron una técnica para extraer y visualizar retículos pigmentados mediante operadores morfológicos. Investigaron el espesor y la variabilidad de este en los bordes del retículo y el tamaño. Los resultados que obtuvieron utilizando operadores morfológicos fueron puramente cualitativos.

Fischer et al. [31] usaron ecualización local del histograma y operadores morfológicos de nivel de gris para realzar los patrones reticulares.

Anantha et al. [32] en el año 2004 compararon dos métodos: el primero se trataba de un algoritmo que se basaba en el filtrado con máscaras de energía de Laws, las cuales son máscaras bidimensionales derivadas de tres filtros unidimensionales simples, utilizando diversas máscaras y obteniendo el cuadrado de las respuestas. Se logró un buen resultado gracias a una media ponderada de dos máscaras de Laws, cuyos pesos

Capítulo 3. Estado del arte

se determinaron empíricamente. Con este sistema se logró un índice de acierto del 78%. Para el segundo método se utilizó la matriz de dependencia vecinal de nivel de gris (NGLDM), sin embargo, no se obtuvieron tan buenos resultados como para el primero.

En el año 2006, Grana, Pellacani y otros [33] propusieron otro enfoque: el uso de núcleos derivativos Gaussianos y el análisis lineal discriminante de Fisher. Se reportó un índice de éxito del 88.3%. Sin embargo, se detectó la existencia de falsos positivos por pelo y por el ruido.

Ese mismo año, Betta et al. [34] comenzaron tomando la diferencia de una imagen con su repuesta a un filtro de mediana. Tras este proceso se creó una máscara aplicando un umbral al resultado de la operación, para así después eliminar las discontinuidades gracias a una operación morfológica. Esta máscara se combinó con una creada a partir de un filtro paso alto aplicado en el dominio de Fourier para discriminar cualquier frecuencia baja moduladora. Los resultados mostraron una sensibilidad del 50% y una especificidad del 100%.

Di Leo et al. [35] en 2008 extendieron este último método y extrajeron características sobre los retículos de la lesión. Los resultados, del 85%, fueron obtenidos mediante el uso de un árbol de decisión, C4.5, para la etapa de clasificación.

Ya en el año 2010, Shrestha et al. [36] presentaron un estudio con un conjunto de 106 imágenes previamente segmentadas para localizar retículos atípicos pigmentados (APN). Si no existía APN, se localizaba de forma manual las regiones con textura irregular. Tras varios intentos y usos de diferentes métricas y clasificadores obtuvieron un resultado del 95.4%.

Wighton [37] publicó un año más tarde, un método basado en el aprendizaje máquina para el análisis de una base de datos compuesta por 734 imágenes clasificadas como Ausente/Presente. Primero, se crea un grupo de 20 imágenes, donde el patrón reticular está presente a lo largo de toda la lesión y otro grupo de otras 20 imágenes. A los píxeles se les asigna una etiqueta: background, ausente, presente. Los píxeles que se encuentran fuera de la región segmentada son etiquetados como background, mientras que los píxeles que están dentro son etiquetados como ausente o presente.

Barata et al. [38] presentaron un método basado en el uso de filtros direccionales y fue aplicado sobre un conjunto de 200 imágenes dermatoscópicas. Se obtuvo un resultado del 86.2% para la clasificación entre Ausente o Presente.

Como se expuso anteriormente en el apartado de análisis de patrones, el patrón reticulado es con el que vamos a trabajar. Para ello, nos basamos en el artículo “Graph-based Pigment Network Detection in Skin Images”, de M. Sadegui et al. [39]. A partir de sus características propuestas lo evaluaremos con otros clasificadores y veremos su precisión.

En su trabajo se propone un método basado en el análisis de grafos para determinar la presencia de patrones reticulares. Es decir, tras un procesamiento previo de la imagen de la lesión y la extracción de sus componentes conectados cerrados, se convierten éstos a grafos para buscar ciclos en los mismos. Estas estructuras cíclicas corresponden a los retículos de la red pigmentada. Para la clasificación se utilizaron 20 características estructurales, 2 geométricas, 37 cromáticas y 10 para la textura. Se obtuvieron finalmente unos resultados del 82% para la clasificación en tres clases (Ausente, Típico y Atípico) y un 93% para dos clases (Presente y Ausente).

En la Tabla 5 se resumen los resultados de clasificación de las obras mencionadas en esta sección.

Algoritmo	Año	Clasificación	Precisión (%)	Imágenes
Anatha et al.	2004	Ausente/Presente	78	155
Grana et al.	2006	No/Parcial/Completo	88.3	60
Betta et al.	2006	Atípico/Típico	-	30
Di Leo et al.	2008	Atípico/Típico	85	173
Shrestha et al.	2010	Melanoma/no	95.4	106
Sadeghi et al.	2010	Ausente/Presente	93	436
Sadeghi et al.	2010	Ausente/Típico/Atípico	82	436
Wrighton et al.	2011	Ausente/Presente	-	734
Barata et al.	2012	Ausente/Presente	86.2	200

Tabla 5. Resultados de la detección de retículo pigmentado

3.2 Otros patrones

Los puntos y los glóbulos son estructuras redondas u ovaladas, de tamaño variable negro, marrón o gris. Es una estructura dermatoscópica que es difícil de distinguir de la red de pigmentos. Esta podría ser la razón por la cual hay tan pocos trabajos en la literatura centrada en su identificación. Yoshino et al. [40] presentaron un algoritmo que utilizaba la operación de cierre morfológico para detectar puntos y glóbulos en la lesión. La operación de cierre usaba un elemento estructural lineal y, posteriormente, se aplicaba un umbral para detectar puntos.

Mirzaalian et al. [41] obtuvieron unos resultados del 91% para la detección de lesiones con patrones streaks. Para ello, primero enfocaron las imágenes para hacer resaltar los detalles, y a continuación, se definió un campo vectorial para cuantificar la dirección radial de las estructuras. Finalmente, se implementó el clasificador usando una máquina de vectores soporte (SVM).

Sadeghi et al. [42] propusieron un algoritmo para la detección de streaks en lesiones pigmentadas. Estas eran clasificadas como ausentes, regulares e irregulares. El objetivo del trabajo era identificar streaks válidos entre un conjunto de líneas que podían ser pelos o simples líneas de la piel.

En el año 2011, Arroyo et al. [43] desarrollaron técnicas supervisadas de aprendizaje automático para detectar patrones con velo azul-blanquecino. Para este fin, se extrajeron características de color de cada píxel individualmente y el clasificador utilizado fue el algoritmo C4.5, que generó el árbol de decisión. Las áreas candidatas seleccionadas en el paso anterior fueron posteriormente clasificadas como melanoma con velo azul-blanco. El método alcanzó una sensibilidad de 80.50% y una especificidad de 90.93%.

En cuanto a patrones hipo pigmentados, Dalal et al. [44] expusieron un método para discriminar melanomas de lunares benignos mediante la detección automática de áreas blancas y las características de medición de estas áreas blancas. Con el fin de identificar áreas blancas e hipo pigmentadas, se determinaron umbrales para cada color basándose en el análisis de histograma de color utilizando un conjunto de imágenes de entrenamiento.

3.3 Múltiples patrones

Tanaka et al. [45] presentaron un análisis de textura para clasificar un patrón en tres categorías: homogénea, globular y reticular. El área de la lesión se dividió en pequeñas regiones y para cada una de ellas se calcularon 110 características de textura. Estas características se basaron en información de histogramas de intensidad, características estadísticas diferenciales, espectro de potencia de Fourier, matriz de co-ocurrencia y componentes conectados. Como resultado, los patrones podían ser clasificados correctamente en tres categorías en un 94% de los casos.

Serrano y Acha [46] plantearon una nueva metodología basada en el reconocimiento de patrones. La lesión es clasificada en cinco tipos distintos de patrones (reticular, globular, empedrado, homogéneo y paralelo).

Para analizar la textura del mismo, se modela la imagen como un modelo de Markov finito simétrico condicional (finite symmetric conditional Markov, FSCM), para después calcular un vector característico con los parámetros del modelo. Para distinguir los distintos tipos detectados, se aplica el criterio de máxima similitud para maximizar la distribución de probabilidad condicional. Para el patrón reticular se alcanzó un índice de acierto del 90%.

Gola et al. [47] trabajaron en un método, en conjunto con la regla ABCD, que intentaba detectar 3 patrones globales. Para ello, desarrollaron 3 algoritmos. Los dos primeros se basaban en la detección de bordes y en el uso de operadores morfológicos para detectar patrones reticulares y globulares. En el tercero, realizaron análisis de color en el espacio RGB con el objetivo de detectar el patrón de velo azul-blanquecino. Los algoritmos propuestos produjeron una precisión media superior al 85%.

Capítulo 4

Diseño de la solución técnica

El objetivo de este capítulo es exponer la metodología seguida en el presente Trabajo Fin de Grado. Para ello se explican las diferentes fases, mostradas en el esquema de la Figura 14: captura de las imágenes, pre-procesado, búsqueda de ciclos, extracción de características referentes al color, estructura, textura y geometría, y clasificación. El propósito será la detección de la red pigmentada en lesiones de piel que, como ya se comentó previamente, es uno de los patrones claves para distinguir el tipo de lesión presente. Para ello clasificaremos las imágenes en dos clases: red pigmentada ausente o presente. Y si es presente veremos si el retículo es atípico o típico.

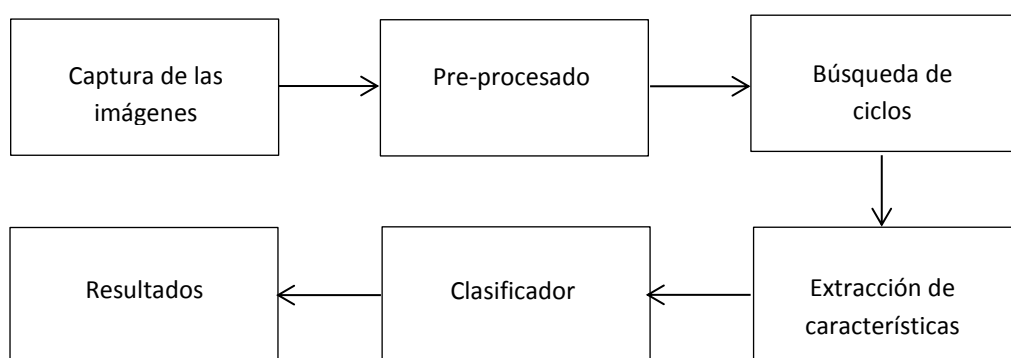


Figura 14. Esquema del sistema

4.1 Captura de las imágenes

Para poder estudiar la eficacia del algoritmo desarrollado nos hemos servido de la base de datos Interactive Atlas of Dermoscopy [49] de la cual se han tomado 724 imágenes dermatoscópicas en color con una resolución de 768×512 píxeles.

A continuación se realiza la primera fase del desarrollo, el pre-procesado.

4.2 Pre-procesado

Esta fase es primordial en cualquier sistema de clasificación, ya que antes de la extracción de las características, se deben dar algunos pasos previos con el objetivo de prevenir el análisis innecesario de los píxeles pertenecientes a la piel que no forman parte de la lesión, así como evitar la influencia de artefactos como burbujas, pelos, etc.

4.2.1 Realce de bordes

La mayoría de las imágenes digitales necesitan corregir su nitidez por lo que, en primer lugar, es necesario enfocarlas mediante un filtro paso alto bidimensional creado a partir de la Ecuación 1. Este filtro paso alto enfoca la imagen eliminando el ruido de baja frecuencia. Se ha tomado α como 0,5.

$$SharpeningFilter(\alpha) = \left(\frac{1}{\alpha + 1} \right) \begin{vmatrix} -\alpha & \alpha - 1 & -\alpha \\ \alpha - 1 & \alpha + 5 & \alpha - 1 \\ -\alpha & \alpha - 1 & -\alpha \end{vmatrix} \quad (1)$$

Las Figuras 15b y 16b muestran el resultado de dos imágenes tras aplicar este paso sobre sus respectivas imágenes originales, Figuras 15a y 16a.



Figura 15a. Imagen Original



Figura 15b. Imagen enfocada

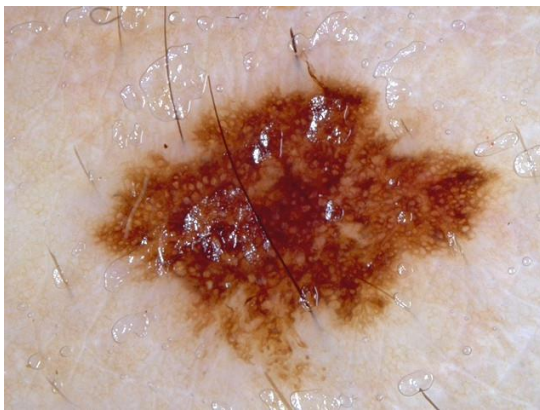


Figura 16a. Imagen Original



Figura 16b. Imagen Enfocada

4.2.2 Eliminación del pelo y las burbujas

Por lo general, las imágenes además de contener la lesión pigmentada de interés, poseen pelos y burbujas. Estas burbujas son el resultado de utilizar medios de inmersión como aceite, alcohol o gel de ecografía que permiten disminuir la reflexión y la refracción irregular en la superficie de la epidermis haciendo que la superficie cutánea se vuelve traslúcida. Sin embargo, tanto las burbujas como el pelo causan ruido que debe ser eliminado para que no perjudique a la hora de la clasificación.

La detección y eliminación del vello es un importante paso que se debe realizar previamente ya que puede ser confundido por el algoritmo como líneas pertenecientes al patrón reticular. Aunque solo llegan a afectar los píxeles que se encuentren dentro de la lesión, ya que el resto son eliminados aplicando una máscara, como mostramos en las Figuras 17 y 18.



Figura 17. Imagen original enmascarada

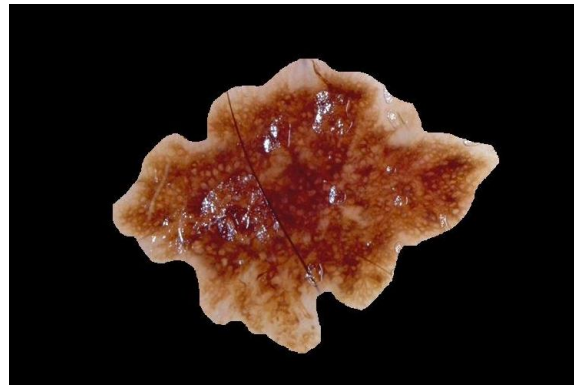


Figura 18. Imagen original enmascarada

El método utilizado para eliminar el vello se basa en buscar y borrar los elementos en los que quepa un elemento estructurante lineal. Este elemento lineal, en nuestro caso, toma diferentes inclinaciones con el fin de buscar aquellos segmentos lineales alargados que puedan aparecer en la lesión pigmentada en cualquier dirección. Una vez encontrados, son eliminados. Mostramos el resultado en la Figura 20.



Figura 19. Imagen original con pelo



Figura 20. Resultado de una imagen sin pelo

4.2.3 Detección de bordes

Los bordes de una imagen digital se pueden definir como transiciones entre dos regiones de niveles de grises significativamente distintos. Aportan una valiosa información sobre las fronteras y permiten identificar las áreas de la región pigmentada.

Una vez que se tiene la imagen limpia y enfocada la detección de bordes es realizada usando el filtro Laplaciano del Gaussiano (LoG), Figura 21. El Laplaciano es una medida bidimensional isotrópica de la segunda derivada espacial de una imagen. Para obtener un buen resultado previamente se aplica un filtro Gaussiano para reducir la sensibilidad al ruido, dando así lugar al filtro LoG.

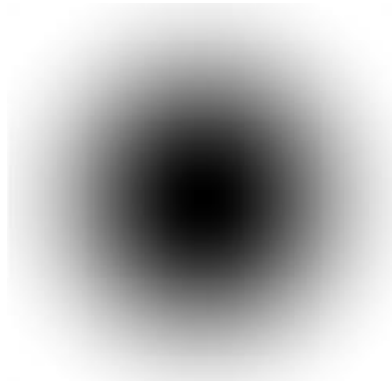


Figura 21. Filtro LoG

El filtro LoG, para el caso de las imágenes de la base de datos, se utiliza para detectar cambios de intensidad del tipo “claro-oscuro-claro” característicos en los segmentos oscuros de los patrones reticulares sobre el fondo de piel más claro. El criterio de detección es la presencia de cruces por cero en la segunda derivada de la imagen. El filtro tiene la forma presentada en la Ecuación 2, para una imagen de entrada $f(x, y)$:

$$h(x, y) = \Delta^2[g(x, y) * f(x, y)] = [\Delta^2 g(x, y)] * f(x, y) \quad (2)$$

$$\Delta^2 g(x, y) = \left(\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right) \frac{-(x^2 + y^2)}{2\sigma^2}$$

donde σ corresponde con la desviación estándar de la ventana gaussiana. Cuanto menor sea su valor mayor será la resolución con la que los bordes son detectados.

A continuación, mostramos en la Figura 23 el resultado de aplicar el filtro LoG.



Figura 22. Imagen original con retículo típico

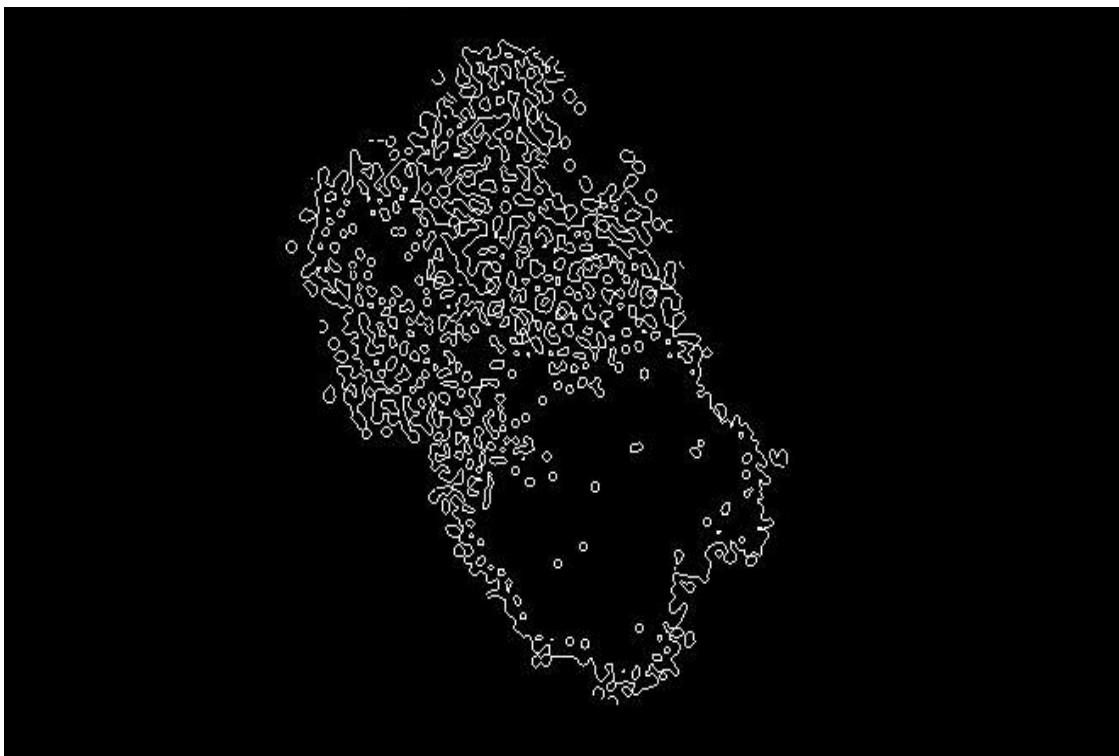


Figura 23. Resultado del filtrado LoG en una imagen con retículo típico

4.3 Grafos

4.3.1 Teoría de grafos

La teoría de grafos es el estudio de estructuras matemáticas utilizadas para modelizar relaciones entre objetos de un conjunto. En este contexto, un grafo $G = (V, E)$ es una pareja ordenada en la que V es un conjunto no vacío de vértices o nodos y E es un conjunto de aristas o enlaces.

Hoy en día, la teoría de grafos es utilizada en numerosos problemas de logística, robótica, genética, sociología, diseño de redes y cálculo de rutas óptimas, entre otros. Son utilizados por ejemplo para modelar trayectos como el de una línea de autobús a través de las calles de una ciudad, en el que se pueden obtener caminos óptimos para el trayecto aplicando diversos algoritmos como puede ser el algoritmo de Floyd. Para la administración de proyectos, utilizando técnicas de revisión y evaluación de programas (PERT) en las que se modelan los mismos utilizando grafos y optimizando los tiempos.

Una importante aplicación de la teoría de grafos es en el campo de la informática, ya que ha servido para la resolución de importantes y complejos algoritmos. Un claro ejemplo es el Algoritmo de Dijkstra, utilizado para la determinación del camino más corto en el recorrido de un grafo con determinados pesos en sus vértices.

4.3.2 Búsqueda de ciclos

Una vez detectados los bordes de la lesión pigmentada mediante todos los pasos anteriores de pre-procesado, pasamos a construir el grafo. Cada píxel detectado como parte de un borde será un nodo del grafo, y todas las conexiones entre ellos conforman las aristas. Para obtener el grafo se obtienen todos los componentes conectados a cada píxel, con conectividad a 8 píxeles.

Ahora ya podemos buscar estructuras cíclicas de la imagen. En trabajos anteriores [30], [33] estas estructuras se localizan utilizando operaciones morfológicas, en concreto secuencias de funciones de apertura y cierre. En el artículo en el que se basa el presente trabajo se decidió buscar los bucles mediante grafos dado que el otro enfoque puede causar numerosos errores.

Capítulo 4. Diseño de la solución técnica

Para realizar la búsqueda de bucles en grafos se necesita calcular, para cada componente encontrado en el grafo su matriz de adyacencia. Esta matriz, con un tamaño $n \times n$ siendo n es el número de nodos, tendrá la siguiente forma:

- $A_{ij} = 1$, si el píxel al que represente el nodo i y el del nodo j son adyacentes.
- $A_{ij} = 0$, en caso contrario.

Para calcular estos valores se toma cada píxel nodo y, tras buscar las coordenadas de todos los elementos no nulos de la matriz, se itera sobre los mismos para comprobar si esos píxeles están conectados, para devolver así la matriz de adyacencia. Después se convierte esta matriz de adyacencia en una lista que indica las conexiones entre nodos. Esta es la forma en la que trabaja el algoritmo ILCA (Iterative Loop Counting Algorithm) [50].

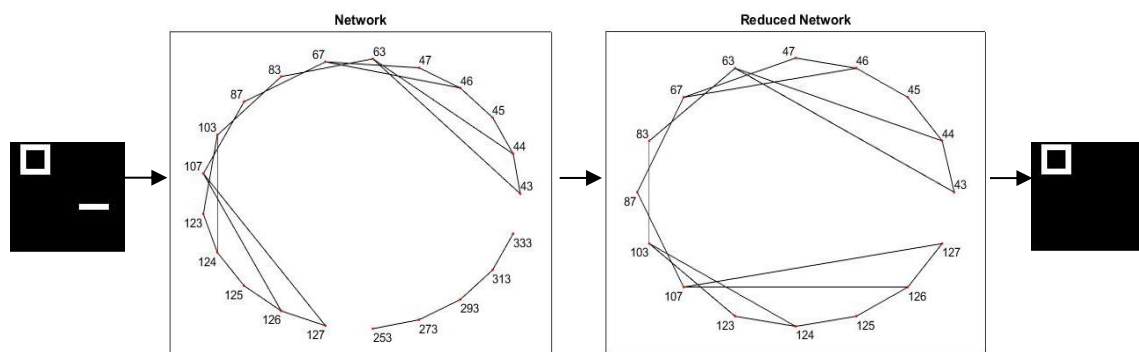


Figura 24. Ejemplo de una imagen resultante de aplicar grafos.

En la Figura 24 primero se obtienen 21 nodos y 24 vértices. Una vez reducido pasa a tener 16 nodos y 20 vértices, dando como resultado la misma imagen que al principio pero solo con el ciclo dibujado en blanco, la raya desaparece al no ser un bucle cerrado.

Una vez ordenados los nodos elige el comienzo más adecuado y comprueba arista por arista si se comparte ésta con el nodo anterior. Si no es así, comprobará si se trata de un bucle (el nodo es el nodo inicial) o no. De forma más detallada, el algoritmo tendrá la siguiente forma:

1. Convertimos la lista de nodos y aristas en una estructura formada por dos campos: "nodes" y "edges". El primero de ellos es un indicador del nodo actual, mientras que el segundo es un vector que representa todos los nodos conectados con el que apunta el campo "node".
2. Calculamos el número de aristas y nodos.

3. Buscamos el nodo inicial óptimo (el que tenga el menor número de aristas) e inicializamos la ruta a estudiar, la primera arista y la lista de bucles.
4. Iniciamos la búsqueda de bucles. Esta búsqueda se realiza estudiando todas las rutas posibles en el grafo hasta que la ruta en cuestión tenga un tamaño menor o igual que uno o que la arista seleccionada esté vacía.

El resultado de la búsqueda de ciclos será una imagen binaria conteniendo todos los bucles detectados, a la que llamaremos net mask, como la que mostramos en la Figura 26.

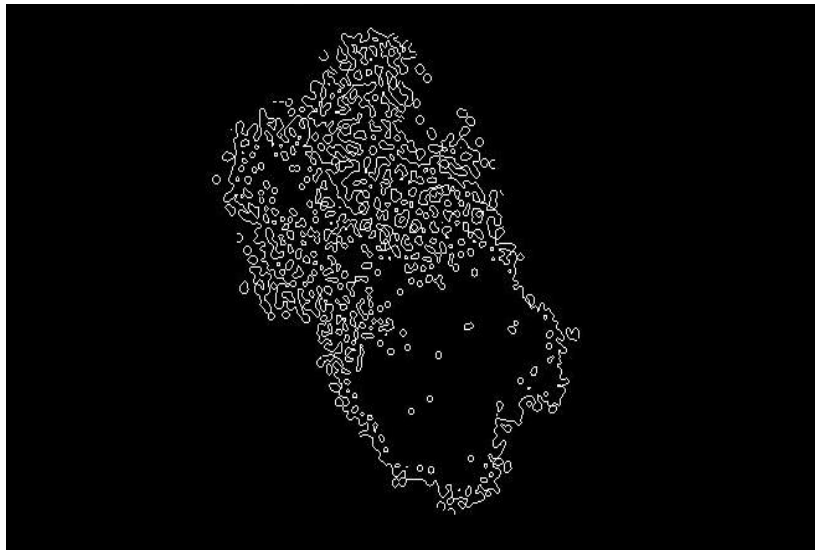


Figura 25. Resultado de aplicar filtro LoG

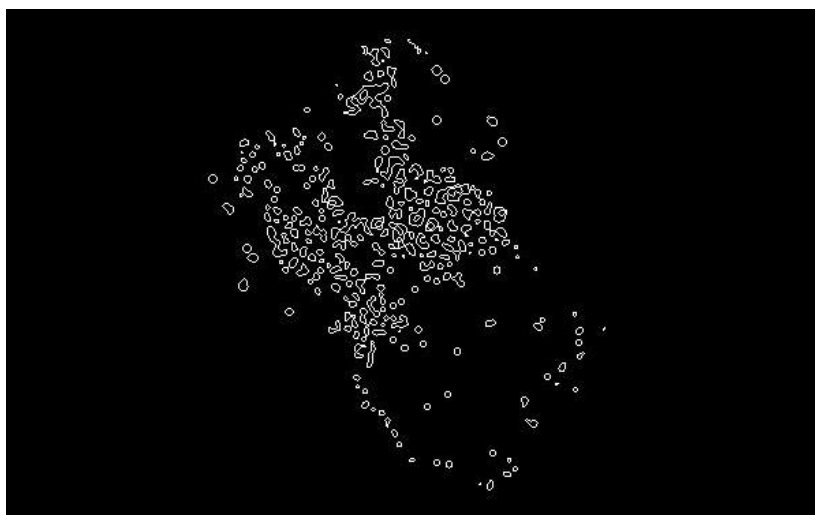


Figura 26. Resultado de aplicar la búsqueda de ciclos mediante grafos

Capítulo 4. Diseño de la solución técnica

Después de encontrar los agujeros que pueden representar los retículos de la red pigmentada, se eliminan las estructuras detectadas erróneamente como pueden ser glóbulos y puntos de acuerdo a los siguientes criterios.

- Fase 1: Para que se trate de un retículo y no de un glóbulo o punto marrón, los agujeros de la red de pigmentos deben tener una intensidad media más alta que el borde; lo contrario es cierto para los glóbulos y puntos marrones.

Además, se eliminan los ciclos encontrados que sean menores de 7 píxeles y mayores de 150 píxeles.

- Fase 2: Las burbujas son estructuras similares a los retículos en términos de intensidad media aunque más brillantes en su interior. Por esto, buscamos en el interior si al menos un píxel tiene una intensidad mayor a 0,8 en una escala de 0 a 1. Si es así, la región entera será considerada una burbuja.

A continuación, mostramos en la Figura 28 el resultado de aplicar el primer paso sobre la imagen obtenida en la búsqueda de ciclos.

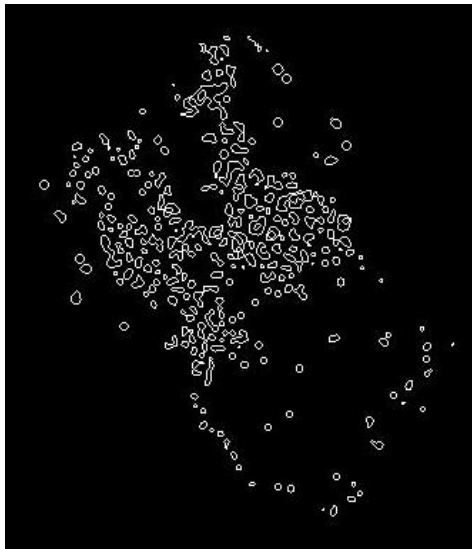


Figura 27. Net Mask

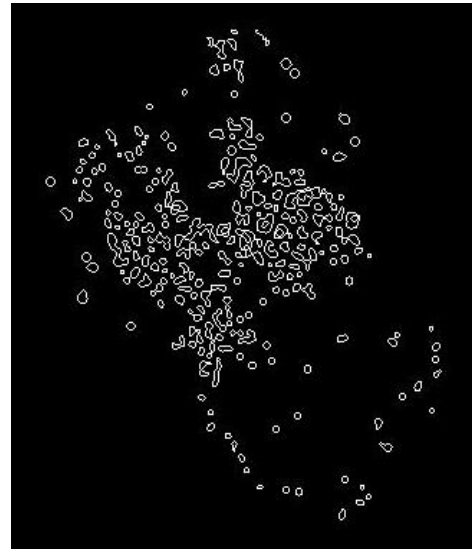


Figura 28. Resultado de aplicar la primera fase

Para visualizarlo con mayor detalle se ha superpuesto este resultado obtenido a la imagen original. El resultado lo mostramos en la Figura 29.

El segundo paso, en el que los quistes y burbujas son identificados, lo mostramos en la Figura 31. Una vez detectados, el algoritmo descarta estos artefactos como posibles agujeros que forman parte del retículo. El resultado final lo mostramos en la Figura 32.

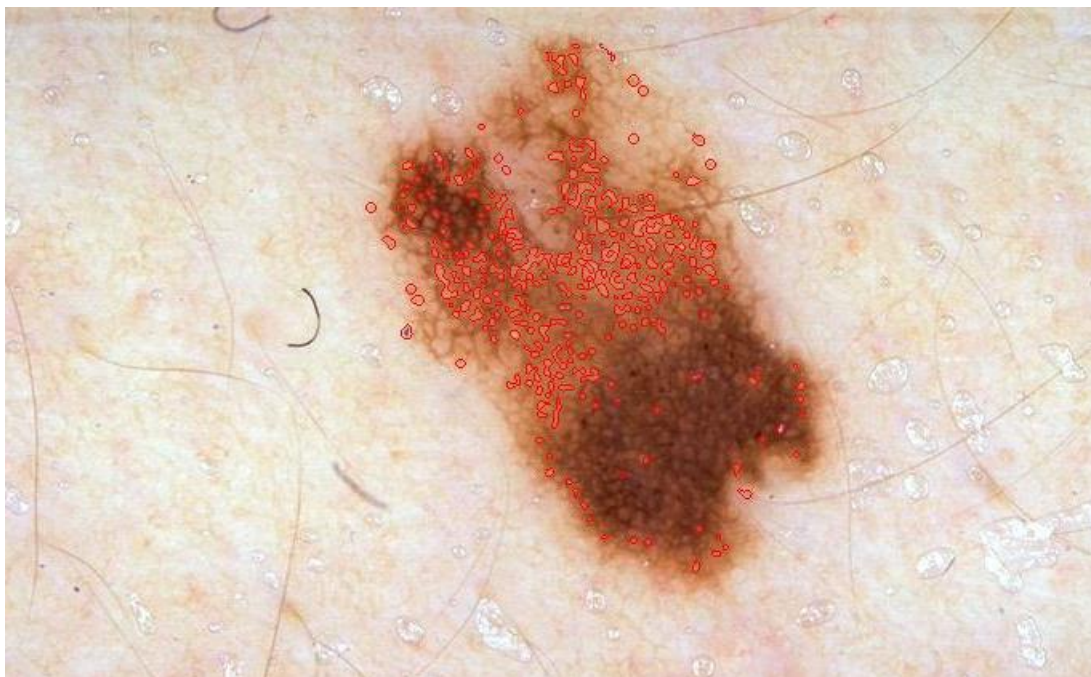


Figura 29. Resultado de la fase 1



Figura 30. Imagen original en blanco y negro

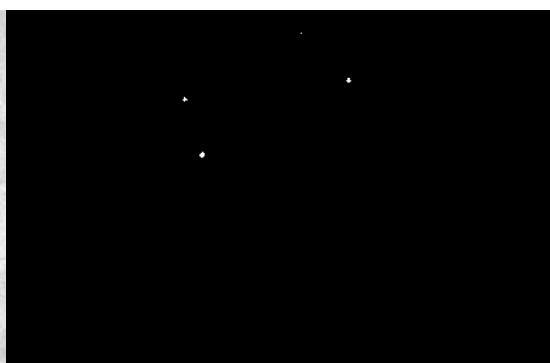


Figura 31. Detección de quistes y burbujas

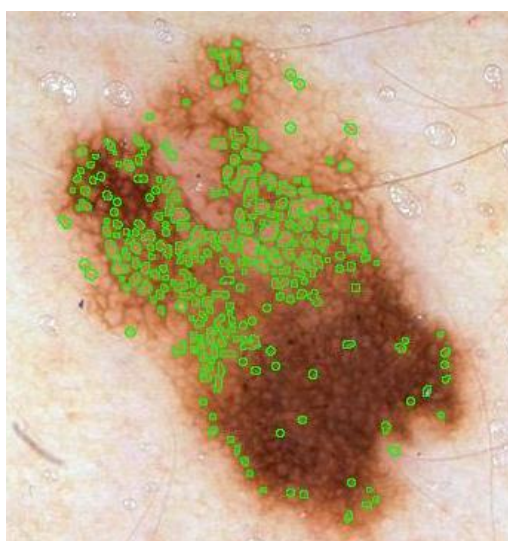


Figura 32. Resultado de la fase 1 y 2

4.4 Extracción de características

La extracción de características hace posible la reducción de la dimensionalidad de una imagen mediante la transformación de los datos de entrada en un conjunto de rasgos que representan las características esenciales de la imagen. Sin embargo, no todas las características son igualmente significativas. Usar un número excesivo de ellas puede incrementar la complejidad del clasificador y degradar el rendimiento del mismo.

Las características que explicaremos durante este apartado se obtienen partiendo de las siguientes imágenes:

- La imagen original sin tratar, obtenida directamente de la base de datos. Se trata de una imagen en color, concretamente en RGB.
- La net mask es la imagen obtenida después de aplicar la búsqueda de ciclos mediante grafos.
- A partir de la net mask se obtiene la pigment network rellenando los agujeros en color blanco.
- Mediante un proceso de erosión sobre la pigment network obtenemos la hole mask quedándonos solo con los agujeros, los cuales representan el interior de los retículos.

A partir de este momento se mencionarán las imágenes con su correspondiente nombre.

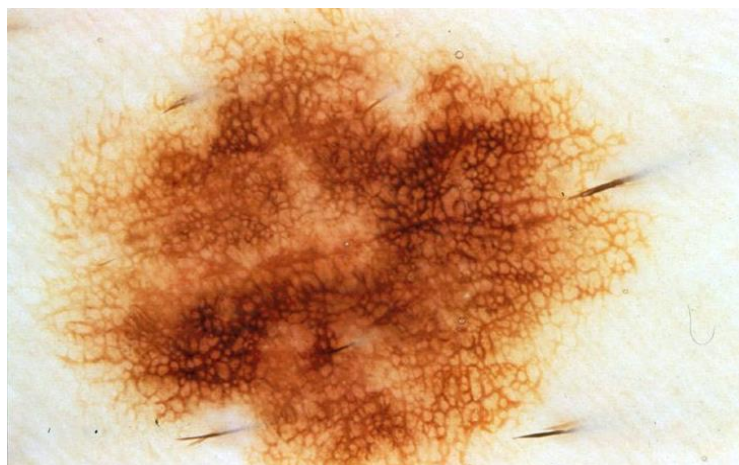


Figura 33. Imagen original de un patrón reticular típico

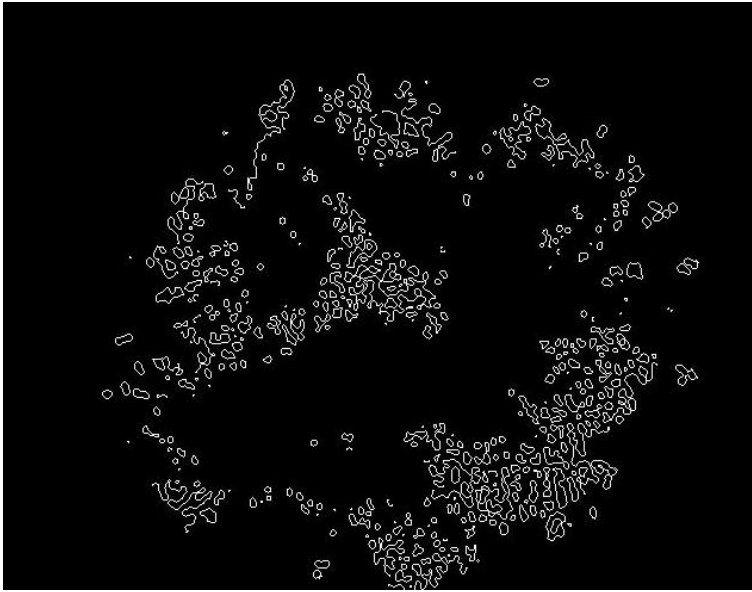


Figura 34. Net mask obtenida tras aplicar la búsqueda de ciclos mediante grafos

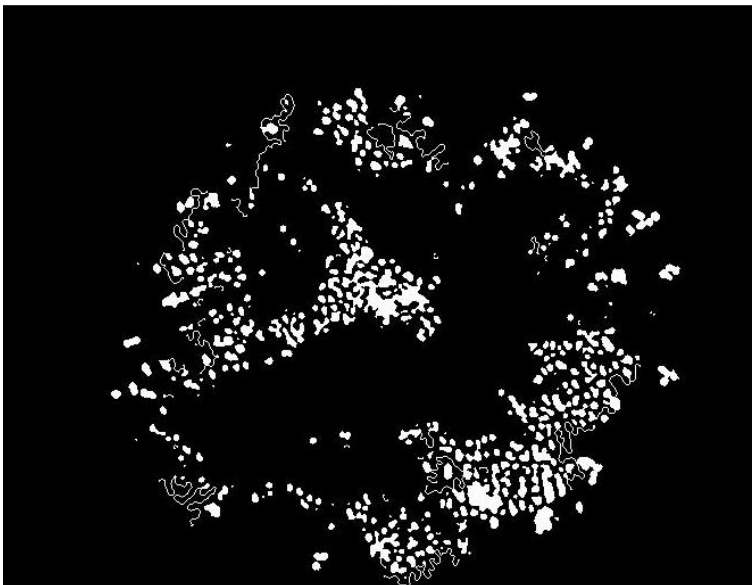


Figura 35. Pigment Network

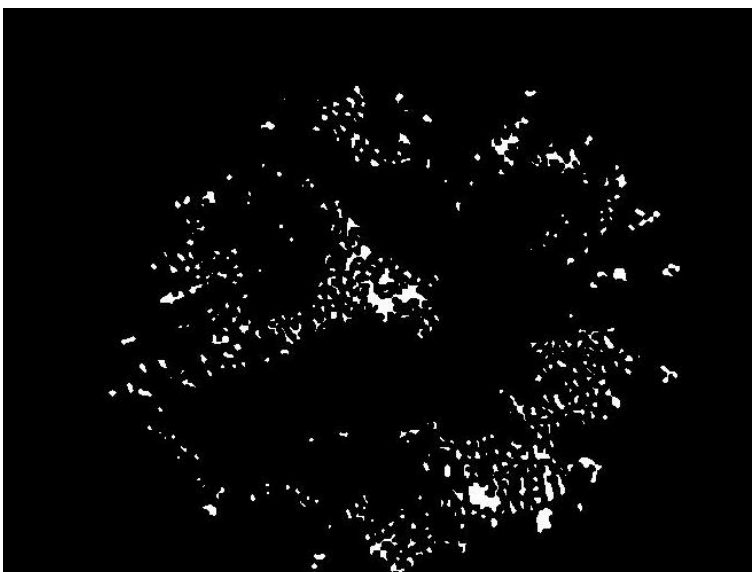


Figura 36. Hole mask

4.4.1 Características de estructura

Entre las características más importantes para el diagnóstico del melanoma se encuentran aquellas que están relacionadas con el tamaño de los retículos. Para analizar la estructura de la lesión se han obtenido un total de 10 características, como mostramos en la Tabla 6. Entre ellas están la media, la desviación estándar y el coeficiente de variación (calculado como la división entre la media y la desviación estándar). Para poder calcular la media y la desviación estándar primero se ha etiquetado cada elemento que se encuentra en la imagen con números consecutivos 1, 2, 3,... teniendo en cuenta la conectividad a 8 píxeles. Después, se ha obtenido un vector de índices. Accediendo a los valores de estos índices la media y la desviación estándar son calculadas tanto para la net mask como para la hole mask.

Como característica importante se ha tenido en cuenta el número de píxeles en la imagen de la net mask y hole mask. Además, para la hole mask se ha calculado el número de holes y un ratio obtenido como la relación del número de píxeles (i,j) blancos de la pigment network entre el número de píxeles blancos en la máscara de la lesión, Ecuación 3.

$$Ratio = \sum_{i,j} \frac{PigmentNetwork(i,j)}{Mask(i,j)} \quad (3)$$

Imagen	Característica
Net mask	Media Desviación estándar Coeficiente de variación Número de pixeles
Hole mask	Media Desviación estándar Coeficiente de variación Ratio

Tabla 6. Características de estructura

4.4.2 Características de geometría

Para determinar la geometría se han utilizado 2 características. La primera de ellas es el ratio de densidad, 'Density Ratio' y se define mediante la Ecuación 4:

$$Density = \frac{|E|}{|V| * \log(LesionSize)} \quad (4)$$

donde $|E|$ es el número de aristas en el grafo G , $|V|$ es el número de vértices o nodos en el grafo y $LesionSize$, es el área de la imagen dentro del límite de la lesión.

Un número menor de nodos para un número fijo de aristas determina que los agujeros del retículo están uniformemente espaciados y el grafo es denso. El tamaño de la lesión se usa para normalizar el cociente; se determinó experimentalmente que lo más apropiado es usar el logaritmo, ya que la relación entre el tamaño de la imagen y su probabilidad para que albergue un patrón reticular no es lineal.

Teniendo en cuenta la definición del patrón se puede deducir entonces que se requiere un ratio de densidad alto para que una imagen sea clasificada como presente.

Además, de 'Density Ratio', se incluye otra característica que es el número de aristas $|E|$ en el grafo.

4.4.3 Características de color

El color juega un papel fundamental en el diagnóstico clínico mediante visión artificial. Por este motivo, primero las imágenes son transformadas del espacio de color RGB al HSV. En el modelo RGB cada color aparece en sus componentes espectrales primarias: rojo, verde y azul. Este modelo está basado en el sistema de coordenadas cartesianas. Sin embargo, en el modelo de color HSV se sigue una representación en coordenadas cilíndricas que es más cercana a la forma en que los humanos perciben los colores y sus propiedades, pues se agrupan las tonalidades de color, lo cual es distinto al caso RGB donde los colores no están necesariamente tan agrupados.

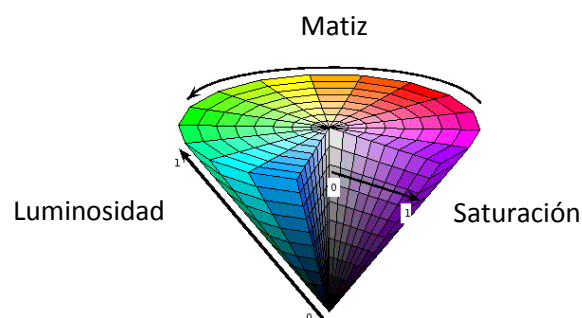


Figura 37. Representación espacio color HSV. Fuente: <http://robocopa.blogspot.com.es/2008/08/bosquejo-en-matlab-del-algoritmo-de.html>

El color HSV está constituido por el matiz, la saturación y la luminosidad, como mostramos en la Figura 37.

El matiz (*Hue*, H) se representa como un grado de ángulo cuyos valores posibles van de 0° a 360°. Cada valor corresponde a un color, por ejemplo: 0° es rojo, 60° es amarillo y 120° es verde. La saturación (*Saturation*, S) se representa como la distancia al eje de brillo negro–blanco. Los valores posibles van del 0 al 100%. A este parámetro también se le suele llamar “pureza” por la analogía con la pureza de excitación y la pureza de color. Cuanto menor sea la saturación de un color, mayor tonalidad grisácea habrá y más decolorado estará. Y por último, la luminosidad (*Value*, V) representa la altura en el eje negro–blanco. Los valores posibles van del 0 al 100%. Dependiendo de la saturación, 100 podría ser blanco o un color más o menos saturado.

Una vez que las imágenes de la hole mask, net mask y la lesion original han sido transformadas al espacio de color HSV se calculan, para cada una de sus componentes, H, S y V, y para la componente verde de RGB: la media, la desviación estándar y el coeficiente de variación (*media/desviación estándar*) en cada una de ellas. En el caso de la media de H, esta ha de ser calculada como la media circular ya que es una magnitud angular.

Adicionalmente a estas 36 características se calcula, solo para la componente verde, una característica llamada ‘atypicality measure’ que corresponde a la suma de los valores de intensidad del canal verde de la imagen net mask normalizado entre el número total de píxeles de coordenadas i,j de esta misma imagen, Ecuación 5.

$$Atypicality = \sum_{i,j} \frac{NetMask_{(:,2)}(i,j)}{NetMask(i,j)} \quad (5)$$

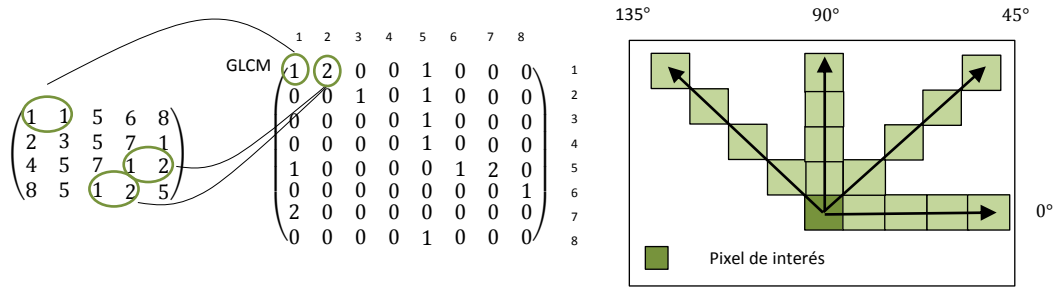
4.4.4 Características de textura

Por textura se entiende la repetición de un patrón cuyos elementos primitivos (texture elements) se organizan siguiendo una estructura determinada. El análisis de la textura intenta aprovechar los rasgos texturales como fineza, rugosidad, granulación, suavidad, aleatoriedad, etc. presentes en la imagen. Para medir la textura de una imagen se utilizan los estadísticos de segundo orden, obtenidos a partir de la matriz de co-ocurrencia de niveles de gris (matriz simétrica).

Haralick et al. [51] desarrollaron un algoritmo, Grey Level Co-occurrence matrix (GLCM), que les permitió discriminar distintos tipos de terrenos en imágenes por

satélite. Posteriormente, se ha aplicado en muchos ámbitos, entre ellos a imágenes biomédicas.

La matriz de co-ocurrencia es una matriz de frecuencias en la que se mide el número de apariciones de un píxel con un nivel de gris i a una distancia específica de otro píxel de nivel de gris j . Mostramos un ejemplo de este tipo de matriz y de cómo se construye en la Figura 38.



Las matrices de co-ocurrencia son medidas de segundo orden porque consideran parejas de píxeles vecinos, separados una distancia d y en un determinado ángulo. Por tanto, las matrices de co-ocurrencia pueden revelar ciertas propiedades sobre la distribución espacial de los grises en la textura de la imagen.

Para este caso, se construyen 2 matrices GLCM, una sobre la imagen de la lesión original y otra sobre la pigment network. Cada una con 4 direcciones: 0° , 45° , 90° y 135° , y con una distancia de 4 píxeles usando 8 niveles de grises promediados para obtener una sola matriz GLCM.

Sobre estas matrices se calculan las siguientes cinco características:

- Entropía:

Este descriptor representado en la Ecuación 6 nos mide la complejidad de la imagen. A mayor entropía mayor complejidad.

$$H = - \sum_i \sum_j p(i,j) \log_2(p(i,j)) \quad (6)$$

- Energía:

$$E = \sum_i \sum_j p(i, j)^2 \quad (7)$$

Esta medida da valores altos cuando en la matriz de co-ocurrencia tiene pocas entradas de gran magnitud, y es baja cuando todas las entradas son similares. Si todos los píxeles son iguales, la Ecuación 7 que representa la energía será mínima. Es una medida de la homogeneidad local.

- Contraste:

$$CON = \sum_i \sum_j (i - j)^2 p(i, j) \quad (8)$$

El contraste cuya fórmula corresponde a la Ecuación 8 representa la cantidad de variaciones locales en los tonos de gris de la imagen. Cuando la variación de los tonos de gris es elevada, el contraste es mayor. Si el contraste es 0, los niveles de gris son constantes a lo largo de la imagen.

- Correlación:

$$C = \sum_i \sum_j \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} p(i, j)$$

$$\mu_x = \sum_i \sum_j ip(i, j) \quad \mu_y = \sum_i \sum_j jp(i, j) \quad (9)$$

$$\sigma_x = \sqrt{\sum_i \sum_j (i - \mu_x)^2 p(i, j)} \quad \sigma_y = \sqrt{\sum_i \sum_j (j - \mu_y)^2 p(i, j)}$$

Es la dependencia lineal de los tonos de gris en la imagen. Algunas propiedades de la correlación son:

1. Un objeto tiene más alta correlación dentro de él mismo que entre objetos adyacentes.

2. Píxeles cercanos están más correlacionados entre sí que los píxeles más distantes.

Si el resultado de la correlación C , Ecuación 9, es 0 entonces no existe relación lineal entre los niveles de gris.

- Homogeneidad:

Es lo opuesto al contraste, por lo que aumenta cuando el contraste entre las parejas de píxeles desciende. La homogeneidad es alta cuando la matriz de co-ocurrencia se concentra a lo largo de la diagonal. Esto ocurre cuando la imagen es localmente homogénea. Mostramos a continuación en la Ecuación 10 la fórmula de la homogeneidad.

$$ISM = \sum_i \sum_j \frac{p(i, j)}{(1 + |i - j|)} \quad (10)$$

En total han sido obtenidas 59 características para cada una de las imágenes de nuestra base de datos.

Capítulo 5

Clasificación y evaluación

La clasificación consiste en predecir el valor de una muestra (clase) basándose en los valores de otras (predictores). Independientemente del tipo de clasificador seleccionado, el proceso de clasificación consta de una serie de pasos:

- Primero se reúnen muestras de objetos de clases conocidas. Se elige un juego de características (vector/matriz de características) apropiado y se calculan las características de los objetos de muestra.
- Después, el conjunto de vectores de características se usa para entrenar el clasificador y se calculan las fronteras entre las clases.
- Se extraen las mismas características de los objetos desconocidos a clasificar.
- Finalmente, el clasificador usa las fronteras calculadas durante el entrenamiento para decidir a qué clases pertenecen los vectores de características de los objetos que queremos reconocer.

Para la clasificación se ha empleado el lenguaje de programación Python (versión 2.7) y la librería de Machine Learning scikit-learn. Como entorno de programación se ha usado Spyder.

5.1 Máquina de vectores de soporte SVM

Las Máquinas de Vectores de Soporte o Support Vector Machines (SVM) son un conjunto de algoritmos de aprendizaje supervisado. Dado un conjunto de entrenamiento de N puntos,

$$(\vec{x}_1, y_1), \dots, (\vec{x}_N, y_N) \quad (11)$$

donde y_i son las etiquetas de las diferentes clases a las que \vec{x}_i puede pertenecer, podemos etiquetar los ejemplos y entrenar una SVM para construir un modelo que prediga la clase de una nueva muestra o conjunto de test.

El clasificador SVM construye un hiperplano que separa el grupo de puntos \vec{x}_i de ambas clases, de manera que el margen entre ellas sea máximo. Suponiendo que es posible separar este conjunto de entrenamiento mediante un hiperplano lineal, los puntos \vec{x}_i que pertenecen al hiperplano satisfacen

$$\vec{w} \cdot \vec{x} + b = 0 \quad (12)$$

donde \vec{w} es un vector normal al hiperplano, y:

$$\frac{|b|}{\|\vec{w}\|} \quad (13)$$

es la distancia perpendicular del hiperplano al origen (se toma $\|\cdot\|$ como la norma euclídea). De entre todos los hiperplanos capaces de separar los datos, como mostramos en la Figura 39, existe un único hiperplano óptimo en el sentido de que es capaz de separar los puntos con el mayor margen de separación entre cada elemento del conjunto de entrenamiento y el hiperplano.

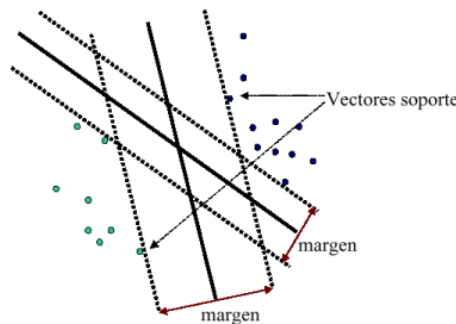


Figura 39. Distintas posibilidades para obtener el margen del hiperplano

Los vectores de características de las dos clases que tengan exactamente la menor distancia al hiperplano se denominan vectores de soporte, Figura 39.

Además, la SVM permite la separación de datos no linealmente separables, Figura 40a, transformando los datos de entrada a un espacio de mayor dimensión conocido como espacio de características en el que sí pueden ser separados mediante un hiperplano, Figura 40c. La transformación se realiza mediante unas funciones matemáticas denominadas núcleo o kernels.

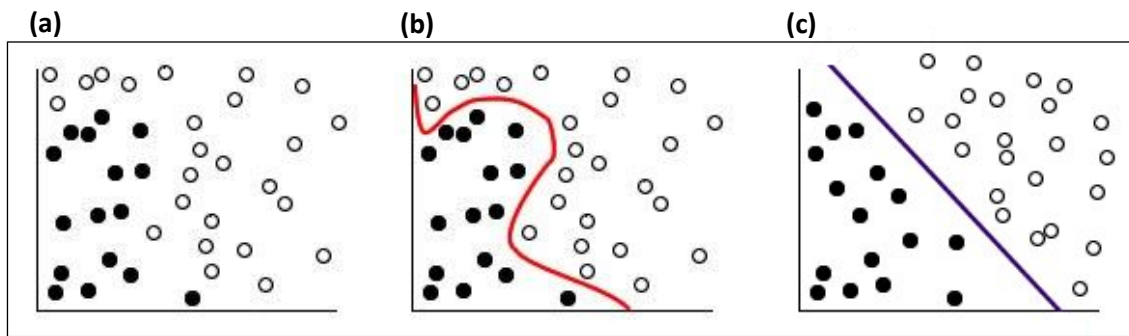


Figura 40. (a) Conjunto de datos original, (b) Datos con separador añadido, (c) Datos transformados

Matemáticamente la función kernel se define como mostramos en la Ecuación 14:

$$(x_i, x_j) \equiv [\phi(x_i) \cdot \phi(x_j)] \quad (14)$$

Siendo las siguientes las más comunes:

- Lineal: $K(x_i, x_j) = x_i \cdot x_j$ (15)

- Gaussiano (RBF): $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ (16)

- Polinómico (heterogéneo): $K(x_i, x_j) = (x_i^T x_j + 1)^n$ (17)

En conclusión, las SVM son una de las técnicas más poderosas del aprendizaje automático. Un punto a favor de utilizar este tipo de modelos es que el desempeño de las SVM no depende del tamaño de la muestra que se va utilizar para el problema, por lo que puede ser utilizado para una cantidad limitada de datos en contraste con otras metodologías que presentan mejor desempeño cuando el tamaño de la muestra es grande. Asimismo, el algoritmo detrás de las SVM se puede ajustar a problemas no

lineales y la solución se realiza bajo programación cuadrática, lo cual hace que su solución sea única y generalizable.

5.2 Evaluación del clasificador

5.2.1 Matriz de confusión

Es una herramienta de visualización que se emplea en aprendizaje supervisado. Cada columna de la matriz representa el número de predicciones de cada clase, mientras que cada fila representa a las instancias en la clase real, como mostramos en la Tabla 7.

En un hipotético caso de clasificador perfecto, todos los elementos de la matriz que no formen parte de la diagonal deberían ser iguales a 0.

		PREDICIÓN	
		Positivo	Negativo
CLASE REAL	Positivo	TP	FN
	Negativo	FP	TN

Tabla 7. Matriz de confusión

La evaluación de la eficiencia del sistema puede ser medida mediante las siguientes métricas:

- **Precision:** Es el número de muestras clasificadas correctamente como positivas respecto al número total de muestras que se detectan como positivas. Se puede entender como la probabilidad de que una muestra detectada como positiva esté bien clasificada.

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

- Recall: Indica la capacidad del clasificador para predecir como casos positivos los realmente positivos. En el caso de una enfermedad, indica la proporción de enfermos correctamente identificados.

$$Recall = \frac{TP}{TP + FN} \quad (19)$$

- Accuracy: Indica la capacidad de un clasificador para predecir los casos correctamente, es decir, pronosticando como positivos los casos realmente positivos y negativos como los realmente negativos. En una enfermedad, el accuracy indica la proporción de casos acertados.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (20)$$

- F1-score: Es la media armónica de la precision y el recall. Cuando las clases presentan una distribución desigual es realmente útil. La fórmula que lo describe se muestra a continuación:

$$F1 = 2 \frac{Precision \cdot Recall}{Precisión + Recall} \quad (21)$$

Debido a que nuestras clases presentan una distribución desigual, teniendo más muestras de unas que de otras, se ha optado por elegir el valor f1 como resultado de nuestro clasificador.

Capítulo 6

Resultados

En este capítulo se presenta la información de la base de datos y los resultados obtenidos en los experimentos realizados utilizando la máquina de vectores de soporte explicada en el capítulo anterior, y la matriz de confusión.

6.1 Descripción del conjunto de datos

En un principio se tomaron 724 imágenes de la base de datos de las cuales 323 presentaban retículo típico, 191 atípico y 210 no tenían retículo, como se muestra en la Tabla 8. Sin embargo, de 83 de ellas no se obtuvieron resultados en la parte de búsqueda de ciclos debido a su coste computacional, por lo que fueron eliminadas, teniendo un total de 641 imágenes dermatoscópicas para usar en clasificación. De este total 291 corresponden a imágenes que presentan retículo pigmentado típico, 170 atípico y las 180 restantes no presentan retículo.

Valor	Número de imágenes	Porcentaje
Atípico	323	44.61 %
Típico	210	29.01 %
Ausente	191	26.38 %

Tabla 8. Tabla de frecuencias de aparición

6.2 Clasificación de las características mediante una SVM lineal

6.2.1 Clasificación de las lesiones con las 59 características obtenidas

En este experimento con el fin de clasificar el patrón reticular entre 2 clases (típico o atípico), han sido eliminadas de la base de datos todas aquellas imágenes en las que éste fuera ausente. Por lo tanto, en este experimento se cuenta con 291 imágenes con retículo típico y 170 con atípico, haciendo un total de 461 imágenes.

El total de las imágenes es dividido en dos conjuntos, uno de entrenamiento (train) y otro de test manteniendo la misma proporción de clases en cada conjunto. En nuestro caso esto es importante debido a que hay más muestras de imágenes con retículo típico que con atípico, es decir, nuestras clases están desbalanceadas.

Posteriormente, hemos calculado el valor óptimo de C (coste) para la SVM. Si damos un valor demasiado grande el modelo penalizará mucho los errores cometidos en el conjunto de entrenamiento y por tanto se producirá overfitting o sobre-entrenamiento, lo que produce que no haya una buena generalización y la clasificación de los nuevos datos de test no sea la correcta. Por otro lado, si C es muy pequeño, el modelo permitirá muchos errores, produciéndose underfitting.

Para evitar el sobre-entrenamiento y obtener una evaluación confiable, se ha aplicado K-fold cross-validation sobre el conjunto de datos de entrenamiento. Esta técnica divide aleatoriamente K veces el conjunto de datos total. Uno de los subconjuntos se utiliza como datos de validación y el resto ($K-1$) como datos de entrenamiento. El proceso de validación cruzada es repetido durante k iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado. En este caso ha sido utilizada la validación cruzada de 10 iteraciones (10-fold cross-validation).

Una vez obtenido el valor de la C óptima es importante normalizar los datos debido a que estos no están definidos en las mismas escalas numéricas y en algunos casos siguen diferentes distribuciones.

Finalmente, se procede a la clasificación mediante una SVM lineal. Esta ha sido elegida en lugar de la polinómica o la gaussiana debido a la cantidad de características extraídas, en total 59.

6.3 Evaluación para la clasificación ausente/presente.

Para este experimento se han tomado de la base de datos tanto las imágenes que presentan retículo atípico como típico considerando ambas como presente y aquellas imágenes que no tienen retículo, es decir, ausentes. En total se han usado, 461 imágenes presentes (típico o atípico) y 180 ausentes, es decir, 641 imágenes.

Primero se ha realizado la clasificación utilizando las 59 características extraídas de la misma forma que para la clasificación típico o atípico y después se ha utilizado únicamente la característica geométrica 'Density Ratio' como se hace en el trabajo [39] en el que este proyecto está basado. Si esta característica toma valores altos quiere decir que la lesión posee retículo pigmentado, es presente, mientras que si es bajo la lesión será clasificada como ausente.

6.4 Resultado de la evaluación del clasificador

Teniendo en cuenta todas las características extraídas sobre sobre 461 imágenes de las que 291 presentan retículo típico y 170 se han obtenido los resultados mostrados en la Tabla 9.

Métricas	Empleando todas las características
Precision	76.5 %
Recall	77.5 %
Accuracy	77%
F1-score	77%

Tabla 9. Resultados clasificación retículo típico o atípico

En algunas imágenes no es fácil distinguir si el patrón reticular es atípico incluso por los expertos. La Figura 41 muestra una de estas imágenes donde el patrón reticular es clasificado como típico por el algoritmo siendo realmente atípica. Un 82.5% de dermatólogos la clasificaron como atípica, y el 17.5% como típica.



Figura 41. Patrón reticular atípico clasificado como típico

Los resultados obtenidos en la clasificación entre ausente o presente se muestran en la Tabla 10. La primera columna corresponde a los valores de las métricas para el caso de tomar las 59 características, la segunda columna para el caso de usar la característica 'Density Ratio' únicamente y la tercera para una selección de características. En todos ellos se han utilizado 641 imágenes de las cuales 180 no tienen retículo (ausente) y 461 sí (presente).

Métricas	Empleando todas las características	'Density Ratio'	Selección de características
Precision	72 %	78 %	80 %
Recall	73 %	73 %	82 %
Accuracy	73 %	73 %	80 %
F1-score	73 %	75 %	81 %

Tabla 10. Resultados clasificación ausente o presente

Estos resultados difieren de los que en el trabajo [39] se presentan. En él para la evaluación ausente/presente utilizando la característica 'Density Ratio' los resultados obtenidos son del 94.3% (Accuracy). Esto es debido a que realizan una selección de imágenes de la base de datos Interactive Atlas of Dermoscopy [49] en la que las imágenes con pelos y burbujas son eliminadas además de que la selección de estas imágenes se realiza de forma proporcional, es decir, se eligen 500 imágenes de las cuales la mitad es clasificada como ausente y la otra mitad como presente. En nuestro

trabajo no se realiza ni la selección de imágenes ni estas están distribuidas proporcionalmente como mostramos en la Tabla 8. Aparte de esto para la obtención de los resultados se ha utilizado cross validation, cuando en el trabajo en el que éste está basado no ha sido usado.

Debido a que no todas las características son igual de importantes para la clasificación ausente/presente y que en nuestro caso con 'Density Ratio' solo no es suficiente, se ha decidido aplicar el clasificador usando un subconjunto de características. Este subconjunto se ha seleccionado en función de los pesos $|w|$ con mayor valor de la SVM lineal. Como resultado 9 características han sido elegidas obteniendo los resultados que en la Tabla 10 mostramos en la tercera columna "selección de características:

- Característica geométrica: Density Ratio.
- Característica de textura: Contraste de la lesion mask.
- Característica de color: Media de los valores del matiz (H) y el coeficiente de variación de la componente verde (G) de la hole mask, la media de los valores del matiz (H) y la saturación (S) y el coeficiente de variación de la componente verde (G) de la net mask.
- Característica de estructura: Número de píxeles y número de agujeros en la hole mask.

Capítulo 7

Conclusiones y futuras líneas de trabajo

En este proyecto se describe un método de aprendizaje automático para la detección de patrones reticulares que son característicos de los melanomas sin necesidad de tener que realizar al paciente una biopsia. Para alcanzar el objetivo propuesto se han seguido las siguientes etapas: pre-procesamiento, búsqueda de ciclos, extracción de características, en total 59, y clasificación mediante máquinas de vectores de soporte (SVM) con kernel lineal. Para llevarlo a cabo se han utilizado 724 imágenes de lesiones pigmentadas obtenidas del Interactive Atlas of Dermoscopy [\[48\]](#).

La clasificación se ha realizado en base a si la lesión presenta retículo o no, es decir si es ausente o presente, y en el caso de que sea presente si éste es típico o atípico.

El primer clasificador presenta unos resultados bastante razonables, con un valor f1 del 81%, teniendo en cuenta el número de imágenes con las que se cuentan. El segundo método en el que se clasifican retículos típicos y atípicos sus resultados no han sido tan buenos aunque sí esperados, obteniendo un valor f1 de 77%. Este resultado es peor ya que incluso para los propios dermatólogos es complicada la diferenciación entre ambas clases.

Como trabajo futuro, se podría ampliar la base de datos de imágenes, mejorar el proceso de filtrado de tal manera que se permita una eficiente extracción de características que mejore el entrenamiento y evaluación en la clasificación. Durante esta parte, se podría hacer una mejor selección y análisis de las características

Capítulo 7. Conclusiones y futuras líneas de trabajo

calculadas. Aparte, otra futura línea de trabajo podría ser detectar otros patrones, como el globular, velo azul-blانquecino, expuestos en el Capítulo 2... De esta forma, se podría diagnosticar si existe melanoma centrándonos en más de un patrón.

Capítulo 8

Gestión del proyecto

En este capítulo se expone la planificación del Trabajo de Fin de Grado que se ha llevado a cabo y el presupuesto del mismo.

8.1 Planificación del proyecto

En este apartado se hace un desglose de todas las tareas y subtareas abordadas durante la ejecución de este proyecto y se acompaña de un diagrama de Gantt que mostramos en la Figura 42.

La fecha de inicio del proyecto es el 5 septiembre de 2016 al seleccionar como objeto el Trabajo de Fin de Grado “Detección del cáncer de piel mediante técnicas de visión artificial” y la fecha de finalización del mismo es el 25 de septiembre de 2017.

El trabajo podría dividirse en las siguientes tareas:

- Primera reunión: En esta tarea se expusieron las bases del proyecto y es necesaria para comprender a qué problema se le quiere dar solución.
- Fase de planificación: Establecimiento de objetivos y estudio del tema en profundidad, dado que el cáncer de piel no tiene relación con la carrera y por tanto el conocimiento sobre él es muy pequeño.
 - Establecimiento de objetivos a conseguir.

- Selección de herramientas a utilizar para el desarrollo del algoritmo.
 - Estudio del melanoma y del patrón reticular en concreto.
 - Investigación sobre el tratamiento de imagen.
 - Investigación sobre la teoría de grafos.
 - Investigación sobre los métodos de clasificación.
- Fase de desarrollo: En esta tarea se plantea cómo solucionar el problema y está dividida en tres subtarefas. A la vez que se va desarrollando el algoritmo se debe ir evaluando para ver si se obtienen los resultados requeridos y si no es así rectificarlos.
 - Arquitectura del sistema
 - Desarrollo del algoritmo
 - Clasificación y Evaluación
 - Fase de documentación: esta fase se ha ido realizando durante todo el tiempo que el proyecto ha tenido lugar.
 - Redacción de la memoria.
 - Entrega.

8.2 Presupuesto

En este apartado se presenta un presupuesto estimado para el proyecto desde su comienzo hasta su finalización. Los principales elementos se describen a continuación:

- **Coste laboral:** Corresponden al salario de un ingeniero junior durante el tiempo del proyecto. El salario estimado es de 25 €/h. Para calcular las horas de trabajo se ha asumido un día de trabajo de 2 horas y 20 días/mes (excluyendo vacaciones), teniendo en cuenta que el proyecto comenzó el 5 de septiembre de 2016 y terminó el 24 de septiembre de 2017. Estos costes se muestran en la Tabla 11.

	Horas de trabajo	Coste/hora	Total
Coste laboral	500 h	25 €/h	12.500 €

Tabla 11. Coste laboral

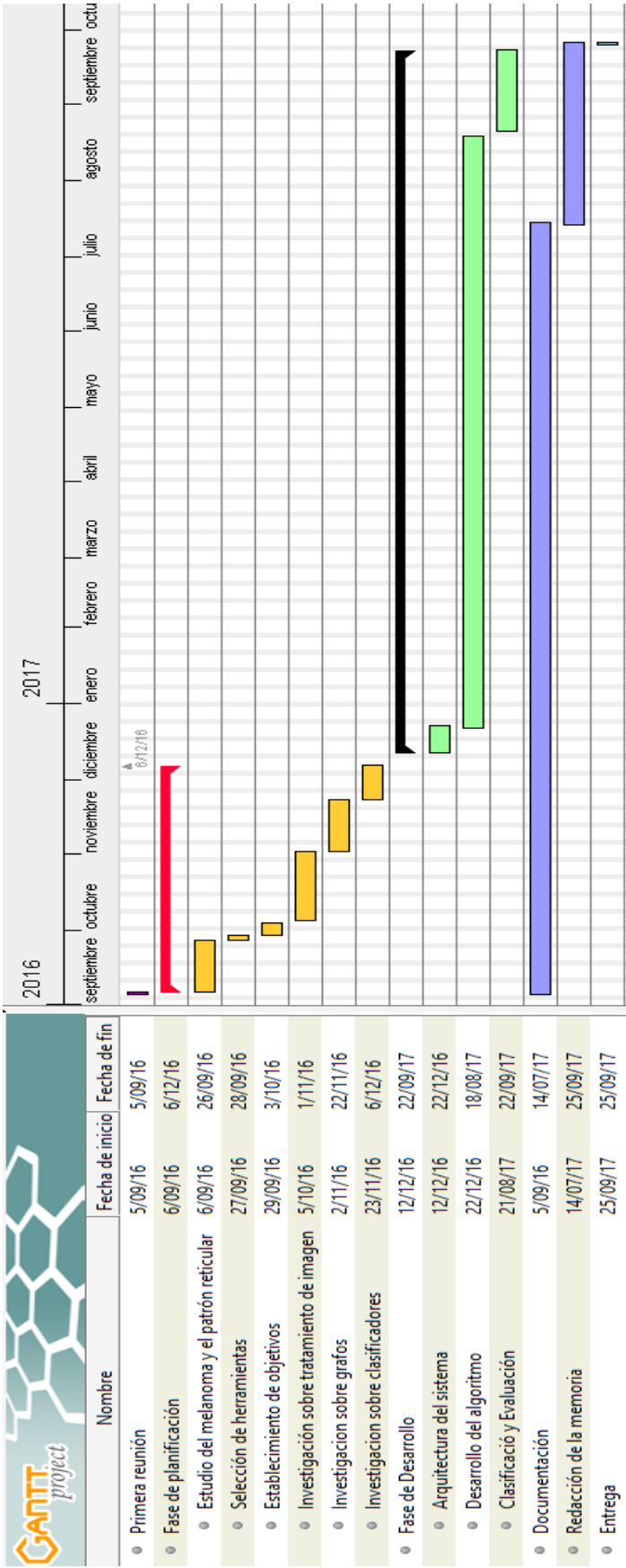


Figura 42. Diagrama de Gantt

- **Coste de equipamiento:** Estos gastos incluyen el coste de amortización del software y hardware necesario durante el proyecto.
 - Software: Para la realización del trabajo se han utilizado MATLAB, para el procesamiento de imágenes y extracción de características y Python para la clasificación. Se ha utilizado la versión académica de MATLAB proporcionada por la universidad, por ello su coste es nulo. Del mismo modo, Python es gratuito siendo su coste también nulo.
 - Hardware: Se ha utilizado un ordenador Toshiba con un procesador Intel® Core™ i7-3630QM CPU 2,40GHz valorado en 650€ con una esperanza de vida de 5 años.

La Tabla 12 muestra un resumen del coste de equipamiento.

	Precio inicial	Vida útil	Uso estimado	Coste de depreciación
Ordenador Toshiba	650 €	6 años	500 h	8 €
			Total	8 €

Tabla 12. Coste de equipamiento

- **Coste computacional:** Dado que a la hora de realizar búsqueda de ciclos en las imágenes dermatoscópicas llevaba mucho tiempo computacional, se optó por el uso de HPCC⁴. Una estimación basada en el centro CESGA español ofrece un precio de 0,2€ por hora computacional y procesador. Se estima que han sido requeridos para todo el proyecto 2 horas usando 1 procesador por imagen, siendo un total de 724 imágenes. Por lo tanto, el coste total es 289 €.
- **Costes inesperados:** con intención de cubrir costes inesperados como puede ser un tiempo de desarrollo que supera al estimado, se ha pensado destinar un 10% del coste de las horas de trabajo, es decir 1.250 €.

Teniendo en cuenta todos los costes expuestos anteriormente, la Tabla 13 muestra el coste total estimado.

⁴ HPCC stands for High-Performance Computing Cluster

Detección del cáncer de piel mediante técnicas de visión artificial

Coste laboral	12.500 €
Coste de equipamiento	8 €
Coste computacional	286 €
Costes inesperados	1.250 €
TOTAL	14.044 €

Tabla 13. Coste total estimado

Bibliografía

1. American Academy of Dermatology, "Cáncer de Piel: Tratable y Derrotable con la Detección Temprana", 2017.
2. Koh, H.K., Miller, D.R., Geller, A.C., Clapp, R.W., Mercer, M.B., Lew, R.A.: Who discovers melanoma? : patterns from a population-based survey. J. Am. Acad. Dermatol. 26(6), 914–919 (1992)
3. Melanoma [Online]
<http://www.lineaysalud.com/enfermedades/melanoma-cancer>
4. Melanoma[Online]
<http://www.uv.es/derma/CLindex/CLmelanoma/CLmelanoma.htm>
5. Cáncer de piel de células basales y células escamosa [Online].
<https://www.cancer.org/es/cancer/cancer-de-piel-de-celulas-basales-y-escamosas/causas-riesgos-prevencion/que-lo-causa.html>
6. Maider Sierra, Ainara Ramírez, "Prevención cáncer de piel 2015", pp.3, Verano 2015
7. Cáncer de piel: las claves de una buena protección solar [Online].
https://elpais.com/elpais/2017/06/12/mamas_papas/1497258932_482399.html
8. Qaisar Abbas,M. Emre Celebi, Irene Fondón, Waqar Ahmad, Melanoma recognition framework based on expert definition of ABCD for dermoscopic images, skin research and technology, 2012 (in press).
9. G. Argenziano, H.P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera et al., Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet, J Am Acad Dermatol, vol. 48, no. 5, pp. 679-93, 2003.
10. La piel [Online]
<http://www.onmeda.es/anatomia/piel.html>
11. Melanoma[Online]
<http://www.cuidateplus.com/enfermedades/cancer/melanoma.html>
12. Cáncer de piel en personas de piel oscura [Online].
<http://www.cancerdepiel.org/prevencion/cancer-de-piel-en-personas-de-piel-oscura>
13. Gandini S,Sera F,CattaruzzaMS, Pasquini P, Picconi O, Boyle P, et al. Meta-analysis of risk factors for cutaneous melanoma II. Sun exposure. Eur J cancer. 2005; 41:45-60.
14. Detectar el melanoma [Online]
http://cadenaser.com/emisora/2017/06/14/radio_valencia/1497436308_122706.html
15. W. Stolz, A. Riemann, A. B. Cognetta, L. Pillet, W. Abmayr, D. Holzel, P. Bilek, F. Nachbar, M. Landthaler, and O. Braun – Falco, "ABCD rule of dermatoscopy: a

- new practical method for early recognition of malignant melanoma," European Journal Dermatology, vol. 4, pp. 521 – 527, 1994.
16. Menzies SW, Ingvar C, McCarthy WH. A sensitivity and specificity analysis of the surface microscopy features of invasive melanoma. *Melanoma Res* 199; 6:55-62.
 17. Argenziano G, Fabroccini G, Carli P, De Giorgi V, Sammarco E, Delfino M. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermoscopy and a new 7-point checklist based on pattern analysis. *Arch Dermatol* 1998; 134: 1563-70.
 18. Salvat. Diccionario Terminológico de Ciencias Médicas. Salvat Editores. Barcelona. 12ª. Ed. 1984
 19. Zaballos P, Carrera C, Puig S, and Malvehy J, Criterios dermatoscópicos para el diagnóstico del melanoma, (2004), 32(1):3–17.
 20. Fases melanoma [Online]
<https://www.aecc.es/SOBREELCANCER/CANCERPORLOCALIZACION/MELANOMA/Paginas/fases.aspx>
 21. Pehamberger H, Binder M, Steiner A, Wolff K. "In vivo epiluminescence microscopy: improvement of early diagnosis of melanoma". *J Invest Dermatol* 1993; 100: 356S-362S.
 22. G. Argenziano, H.P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, et al., Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet, *J. Am. Acad. Dermatol.* 48 (5) (2003) 679–693.
 23. Braun-Falco O, Stolz W, Bilek P, Merkle T, Landthaler M. "Das Dermatoskop. Eine Vereinfachung der Auflichtmikroskopie von pigmentierten Hautveränderungen". *Hautarzt* 1990; 41: 131-136
 24. La dermatoscopia y la fotografía digital, claves en el diagnóstico diferencial y el seguimiento del cáncer de piel [Online]
<http://www.mdanderson.es/noticias/la-dermatoscopia-y-la-fotografia-digital-claves-en-el-diagnostico-diferencial-y-el-seguimie>
 25. Argenziano, G., Soyer, H., De Giorgio, V., Piccolo, D., Carli, P., Delfino, M., Ferrari, A., Hofmann-Wellenhof, R., Massi, D., Mazzocchetti, G., Scalvenzi, M., Wolf, I.H.: Interactive atlas of dermoscopy. EDRA-Medical Publishing and New Media, Milan (2000)
 26. Argenziano, G., Soyer, H.P., Chimenti, S., Talamini, R., Corona, R., Sera, F., Binder, M., Kopf, A.W.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. *J. Am. Acad. Dermatol.* 48(5), 680–693 (2003)
 27. Seidenari S, Pellacani G, Martella A, Giusti F, Argenziano G, Buccini P, et al. "Instrument-, age- and site-dependent variations of dermoscopic patterns of congenital melanocytic naevi: a multicentre study". *Br J Dermatol.* 2006; 155: 56-61.

28. Zalaudek I, Grinschgl S, Argenziano G, Marghoob AA, Blum A, Richtig E, *et al.* "Age-related prevalence of dermoscopy patterns in acquired melanocytic naevi". *Br J Dermatol.* 2006; 154: 299-304.
29. Perednia DA, Gaines JA, Rossum AC. Variability in physician assessment of lesions in cutaneous images and its implications for skin screening and computer-assisted diagnosis. *Arch Dermatol*, 1992, vol 128, pp 357–364.
30. Fleming, M.G., Steger, C., *et al.*: Techniques for a structural analysis of dermoscopic imagery. *Comput. med. imaging graph.* 22(5), 375–389 (1998)
31. Fischer, S., Guillod, J., *et al.*: Analysis of skin lesions with pigmented networks. In: *Proceeding of International Conference Image Processing*, pp. 323–326 (1996)
32. Anantha, M., Moss, R.H., Stoecker, W.V.: Detection of pigment network in dermoscopy images using texture analysis. *Comput. Med. Imaging Graph.* 28(5), 225–234 (2004)
33. Grana, C., Cucchiara, R., Pellacani, G., Seidenari, S.: Line detection and texture characterization of network patterns. In: *Proceedings—International Conference on Pattern Recognition*, vol. 2(1699200), pp. 275–278 (2006)
34. Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., Sommella, P.: Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. *IEEE International Workshop on Medical Measurement and Applications, MeMeA*, art. 1644462, 63–67 (2006)
35. Di Leo, G., Paolillo, A., Sommella, P., Liguori, C.: An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin lesions. In: *Proceedings of IEEE Conference Virtual Environments Human-Computer Interfaces and Measurement Systems*, pp. 190–194 (2008)
36. Shrestha, B., Bishop, J., Kam, K., Chen, X., Moss, R.H., Stoecker, W.V., Umbaugh, S., Stanley, R.J., Celebi, M.E., Marghoob, A.A., Argenziano, G., Soyer, H.P.: Detection of atypical texture features in early malignant melanoma. *Skin Res. Technol.* 16(1), 60–65 (2010)
37. Wighton, P., Lee, T.K., Lui, H., McLean, D.I., Atkins, M.S.: Generalizing common tasks in automated skin lesion diagnosis. *IEEE Trans. Information Technology in Biomed.* 15(4), 5763779, 622–629 (2011)
38. Barata, C., Marques, J. S., Rozeira, J.: A system for the detection of pigment network in dermoscopy images using directional filters. *IEEE Trans. Biomed. Eng.* 59(10), art. no. 100, 2744–2754 (2012)
39. Sadeghi M, Razmara M, Ester M, Lee TK, Atkins MS: Graph-based pigment network detection in skin images. *SPIE Digital Library*. 2010 Apr 26. *Proc. SPIE* 7623 762312-1.
40. Yoshino, S., Tanaka, T., Tanaka, M., Oka, H.: Application of morphology for detection of dots in tumor. In: *Proceedings of the SICE Annual Conference, WPI-3-3*, pp. 407–410 (2004)

41. Mirzaalian, H., Lee, T.K., Hamarneh, G.: Learning features for streak detection in dermoscopic color images using localized radial flux of principal intensity curvature. In: Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis, 6164758, pp. 97–101(2012)
42. Sadeghi, M., Lee, T., Lui, H., McLean, D., Atkins, S.: Detection and analysis of irregular streaks in dermoscopic images of skin lesions. IEEE Trans. Med. Imaging (2013, in press)
43. Arroyo, J. L. G., Zapirain, B. G., Zorrilla, A. M.: Blue-white veil and dark-red patch of pigment pattern recognition in dermoscopic images using machine-learning techniques. In: IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2011, art. no. 6151559, 196–201 (2011)
44. Dalal, A., Moss, R.H., Stanley, R.J., Stoecker, W.V., Gupta, K., Calcara, D.A., Xu, J., Shrestha, B., Drugge, R., Malters, J.M., Perry, L.A.: Concentric decile segmentation of white and hypopigmented areas in dermoscopy images of skin lesions allows discrimination of malignant melanoma. Comput. Med. Imaging Graph. 35(2), 148–154 (2011)
45. Tanaka, T., Torii, S., Kabuta, I., Shimizu, K., Tanaka, M., Oka, H.: Pattern classification of nevus with texture analysis. In: Annual International Conference of the IEEE Engineering in Medicine and Biology—Proceedings 26 II, pp. 1459–1462 (2004)
46. Serrano C. and Acha B., Pattern analysis of dermoscopic images based on markov random fields, Pattern Recogn. (2009), 42(6), 1052–1057.
47. Isasi Gola, A., Garcia Zapirain, B., Mendez Zorrilla, A.: Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. Comput. Biol. Med. 41, 742–755 (2011)
48. Detector automatic de melanomas [Online]
<https://clustersalud.americaeconomia.com/desarrollan-detector-automatico-melanomas>
49. Argenziano, G., Soyer, H.P., De Giorgio, V., Piccolo, D., Carli, P., Delfino, M., Ferrari, A., Hofmann-Wellenhof, V., Massi, D., Mazzocchetti, G., Scalvenzi, M., Wolf, I.H.: Interactive Atlas of Dermoscopy (Book and CD-ROM). Edra medical publishing and new media, Milan (2000)
50. Kirk, J.: Count loops in a Graph. 2007
51. R.M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification", IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-3, pp. 610-621, 1973

Anexo

Summary

The main objective of this project is to develop an algorithm capable of detecting skin cancer through artificial vision techniques. Dermatologists currently use biopsies to find out if a patient has melanoma or not but image-based computer aided diagnosis (CAD) systems, in which is framed our project, have significant potential for screening and early detection of malignant melanoma. Application of computational intelligence methods helps dermatologists in faster data process to give better and more reliable diagnoses.

Every year, millions of people suffer skin cancer [\[1\]](#). Early detection and surgery represent the only treatment for patients who are affected by this disease. Unlike many other cancers, melanoma is visible on the skin but is the most dangerous skin cancer. It represents the 80% of deaths; that is why, it is so important to detect it on time. If melanoma is identified in its first stages (0 and I), it could be cured in 90-95% of cases [\[7\]](#). But if melanoma is not treated the probability of cure is reduced to 45% in the third stage and 10% in the fourth [\[20\]](#).

The skin is the human body's largest organ, with a total area of about 1.5 to 2 m² and a weight of 3.5 kg [\[10\]](#). It serves many important functions, including: regulating body temperature, maintaining water and electrolyte balance, participating in vitamin D synthesis, protecting the body against trauma, etc. Skin is formed of three layers; the first one is the epidermis which provides a waterproof barrier. Beneath the epidermis, is the dermis that contains tough connective tissue, hair follicles, and sweat glands. Finally, the deeper subcutaneous tissue, hypodermis, is made of fat and connective tissue.

Summary

Is in the epidermis where melanocytes, that produce the pigment melanin, are created. When the skin is exposed to the sun for a long time, melanocytes produce more melanin as a defence against the action of ultraviolet rays, UVA and UVB. At the time when melanocytes outnumber the normal skin cells, new moles and freckles appear. This can occur anywhere on the body, but the most commonly affected areas are the back in men and the legs in women [2]. Melanomas are uncommon in areas which are protected from sun exposure, such as the buttocks and the scalp. Nevertheless, several risk factors can make a person more likely to develop melanoma.

- Fair skin, light hair colour, light eye colour: People who are very fair skinned, especially with fair or red hair, are more at risk of developing melanoma. So are people with a lot of freckles. Also, people with darker skins can still get melanoma but they have more natural protection against it [12].
- Age: The risk of melanoma increases with age. So it is more common in older people.
- Family history of melanoma.
- Previous melanoma diagnosis: People who have had a melanoma have an increased risk of getting a second melanoma.
- Exposure to UV radiation or tanning bed use: Sunburn definitely increases the risk of melanoma and research shows that the type of ultraviolet light used in sunbeds (UVA) can cause also all types of skin cancer.

In general, melanoma can be detected in its beginnings. Each person plays an important role in the early detection of skin cancer, when it is more likely to be healed. The clinical diagnosis of melanoma is commonly based on the ABCD system [15], which is a checklist of important parameters. This technique focuses on asymmetry, irregular border, number of colours, and the presence of differential structures (streaks, globular elements, etc.). Another criterion is Menzies method [16], which is based on 11 features divided between negative and positive. The 7-point checklist [17] provides a simplification of standard pattern analysis because of the low number of features, or criteria, to identify. As with the ABCD rule, it can be easily learned and applied and has proven to be reliable in diagnosing melanoma.

Dermatoscopy is a non-invasive technique that allows the visualisation of numerous morphological features that enhance the clinical diagnosis of pigmented skin lesions and can improve the acuteness of the diagnosis of melanoma. With the dermatoscope, which is the instrument used for dermatoscopy, patterns can be observed.

The presence of specific dermoscopic features in different regions of the same lesion contributes to make a diagnosis of melanocytic lesions and is called local patterns. They are dermoscopic structures such as pigment network, dots and globules, streaks, blue-whitish veil, regression structures and hypopigmentation. Global features allow a quick preliminary categorization of a given pigmented skin lesion prior to more detailed assessment, and they are presented as arrangements of textured patterns covering most of the lesion. The main global patterns are reticular, globular, cobblestone, homogeneous, starburst, parallel, multicomponent and unspecific pattern. The most important diagnostic feature of melanocytic lesions is the pigment network, which consists of pigmented network lines and hypo-pigmented holes. These structures show prominent lines, homogeneous or inhomogeneous meshes

Automatic detection of reticular patterns has received a lot of attention in recent years. Anantha et al. [32] in 2004 compared two statistical texture identification methods for detecting the pigment network. The first method was the neighbouring grey-level dependence matrix (NGLDM), and the second method used the lattice aperture waveform set (LAWS). Later, Betta et al. [34] began by taking the difference of an image and its response to a median filter. Di Leo et. al. [35] extended this method and compute features over the 'holes' of the pigment network. A decision tree is learned in order to classify future images and an accuracy of 85% is achieved.

Other patterns such as globular or dots, streaks, Blue–Whitish Veil etc. have been source of interest. Yoshino et al. [39] presented an algorithm that use morphological closing operation to detect dots. The closing operation used a linear structural element. Afterwards, a thresholding is applied to detect dots. Mirzaalian et al. [40] proposed a machine-learning approach to detect streaks which captures the quaternion tubularness in the colour dermoscopic images. In a recent work, Arroyo et al. [42] also proposed supervised machine learning techniques to detect blue-white veil. To this aim, colour features were extracted from each individual pixel and the classifier used was C4.5 algorithm, which generated the decision tree.

SOCIO-ECONOMIC AND LEGAL FRAMEWORK

On the one hand, the data protection law must be complied with. Patients from whom the images have been obtained must be consulted in advance.

On the other hand, computer aid diagnosis reduces the material and hospital personal costs.

Biopsy can cause bleeding, infection, scarring, and accidental injury to adjacent structures. That is why it is best to avoid it so as not to cause the patient any of these damages.

METHODOLOGY

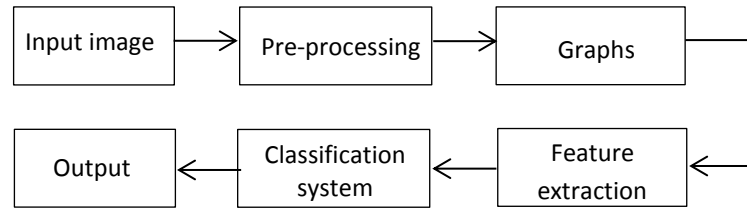


Figura 43. Block diagram representation

- a. **Imagen Acquisition:** In first step image are acquired. We use a set of 724 images obtained from the Interactive Atlas of Dermoscopy [49].
- b. **Image pre-processing:** In second step pre-processing is done using *unsharp* filter to correct the sharpness of the image. After this step hairs and bubbles are removed because they could be interpreted for the algorithm as lines of the reticular pattern, and finally in order to detect changes in the lesion of the skin LoG filter is applied. We used this filter to detect the sharp changes of intensity along the edge of the holes inside the segmented lesion. Because of the inherent properties of the filter, it can detect the “light-dark-light” changes of the intensity well. The result of this edge detection step is a binary image which is subsequently converted into a graph.
- c. **Graphs:** The binary image is converted to a graph (G) using 8-connected neighbours to find holes or cyclic structures of the lesion. Each pixel in the connected component is a node of G and each node has a unique label according to its coordinate. To find round texture features (i.e. holes); all cyclic subgraphs of G are detected using the Iterative Loop Counting Algorithm (ILCA) [50]. This algorithm transforms the network into a tree and does a depth first search on the tree for loops.
- d. **Feature extraction:** In fourth step features are extracting once we have the possible holes detected. These features are related with the shape, spatial, colour and texture of the pigment network. In total, 59 features have been extracted.
 - **Structural features:** The mean, standard deviation (*std*) and coefficient of variation (*mean/std*) have been calculated for two images, hole mask and net mask. Also, we compute the size (number of pixels) for the hole mask and we include the ratio of the network size (number of pixels in the net and the hole mask) to the

lesion size (number of pixels in the segmentation mask). In total ten features are obtained.

- Geometric features: we have defined two features. The first one is called 'Density Ratio' and the second one is the number of edges in the graph G .
 - Chromatic features: colour also plays a crucial role in clinical diagnosis. We therefore convert the image to HSV colour space and compute features over each channel as well as the original green channel of the image. The features are the mean, standard deviation and coefficient of variation (*mean/std*) of the intensity values. Also, 'atypicality measure' is calculated as the sum of the intensity values over the green channel of the pixels in the net mask normalized by the number of the pixels in the net mask.
 - Textural features: in order to measure the texture of an image, the second-order statistics are used, namely a grey level co-occurrence matrix (GLCM). From this matrix, entropy, energy, contrast, correlation and homogeneity are estimated. We construct 2 GLCMs (in the four directions of 0° , 45° , 90° , 135° and within the distance of 4 pixels using 8 gray levels averaged to obtain a single GLCM) and extract the 5 texture metrics from each. The first GLCM is constructed over the entire lesion and the second is constructed over the pigment network.
- e. Classification system: In order to classify the lesions support vector machines (SVM) have been used, which are a discriminative classifier formally defined by a separating hyperplane. In other words, given labelled training data (supervised learning) the algorithm outputs an optimal hyperplane that categorizes new examples. The solution is the hyperplane that gives the largest minimum distance to the training examples and it receives the important name of margin within Support Vector Machine's theory. Therefore, the optimal separating hyperplane maximizes the margin of the training data.
- In addition, SVM allows the separation of non-linearly separable data transforming the input data into a larger dimension space where they can be separated by a hyperplane. The transformation is performed by mathematical functions called kernels. The most popular kernel functions are: linear, Gaussian and polynomial.

- Classification system using linear SVM:
 1. In the first experiment all the features (59) have been taken into account to classify between atypical or typical pigment networks. In total, we have 291 with typical pigment network and 170 with atypical, making a total of 461 images.
 2. In the second experiment, a feature selection was made to see which of them contributed more to the classifier.
 3. Because our project is based on the reticular pattern we have made a classification to differentiate between lesions that present pigment network and not. To achieve this goal we have compared two methods: first using all the features as in previous experiments and second choosing only 'Density Ratio' as feature.

RESULTS

At first, 724 images of the database were taken, of which 323 had typical pigment network, 191 atypical, and 210 did not have. However, for 83 of them we did not obtain results in the search of cycles so they were removed, having a total of 641 dermatoscopic images to use in classification. Of this total 291 correspond to images that are typical, 170 atypical and the remaining 180 are absent.

- Typical or Atypical

The results are shown in Table 14. In the first column we have the different metrics that have been calculated and, in the second one, the results if we choose all the extracted features.

Metrics	59 features
Precision	76.5 %
Recall	77.5 %
Accuracy	77%
F1-score	77%

Tabla 14. Results of typical or atypical classification

In some images, it is not easy to distinguish whether the reticular pattern is atypical even by the experts.

- Absent or Present

The results are shown in Table 15.

Metrics	59 features	'Density Ratio'	Feature selection
Precision	72 %	78 %	80 %
Recall	73 %	73 %	82 %
Accuracy	73 %	73 %	80 %
F1-score	73 %	75 %	81 %

Tabla 15. Results of absent or present classification

In the project in which this work is based the results are different. Using 'Density Ratio' only as a feature, they get 94.3%. This could be because an image selection was carried out, choosing images without bubbles or hair that are the most problematic. Also, in this project a ten-fold cross validation is made during classification and the dataset is not balanced. In [39] they used 250 images that have pigment network and 250 that not.

Because this feature is not sufficient to classify images that are absent from present, it has been decided to make a feature selection.

CONCLUSION AND FUTURE RESEARCH

In this project, we describe an automatic learning method for the detection of reticular patterns that are characteristics in melanomas without having to make a biopsy to the patients. In order to achieve this objective some stages have been performed: pre-processing, search for cycles, feature extraction and classification using linear kernel support vector machines (SVM). In total we have used 724 images of pigmented lesions obtained from the Atlas of Dermatoscopy.

The f1-score of the system is 81 % in classifying images to one of the two classes of absent and present over a dataset consisting of 641 images. The method was also validated on a different set and achieves an f1-score of 77% discriminating between typical or atypical.

Summary

Several tasks can be done to continue with the research line presented in this document. It would be possible to expand the image database, improve the filtering process and it could be made a selection and analysis of the calculated features. Furthermore, another future line of work could be to detect other skin patterns, such as globules and blue-whitish veil.

PROJECT MANAGEMENT

Project planning: The present work began on September 5, 2016 and ended on September 25, 2017. The work could be divided into the following tasks: first meeting, a planning phase where the main objective was exposed and the subject was studied in depth, development phase where the algorithm was made and validate, phase of documentation which has been developing throughout the time the work has lasted, writing of memory and delivery.

- **Project Budget:** Cost sources are split into four main groups: labour costs, cost of equipment (hardware and software), computational time and unexpected costs. Labour costs correspond to the salary of a junior engineer during the timespan of the project. Equipment costs include the amortization cost of the software and hardware needed during the project. Computational time includes the costs associated to the HPC cluster that has been used to perform the simulations and unexpected costs are the 10% of labour costs. Taking into account previous expenses, the estimated total cost of the project is given in Table 16.

Labour costs	12.500 €
Equipment costs	8 €
Computational time	286 €
Unexpected costs	1.250 €
TOTAL	14.044 €

Tabla 16. Estimated budget

